(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number
WO 03/014357 A1

(51) International Patent Classification[7]: C12N 15/60, 15/82, 9/88, 5/10, A01H 5/00, 5/10

(21) International Application Number: PCT/CA02/01051

(22) International Filing Date: 10 July 2002 (10.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/311,282 9 August 2001 (09.08.2001) US

(71) Applicant *(for all designated States except US)*: UNIVERSITY OF SASKATCHEWAN [CA/CA]; Office of Research Services, Room 208 Kirk Hall, 117 Science Place, Saskatoon, Saskatchewan S7N 5C8 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: POZNIAK, Curtis, J. [CA/CA]; 2215 A. St. Patrick Avenue, Saskatoon, Saskatchewan S7M 0L2 (CA). HUCL, Pierre [CA/CA]; 1031 Aird Street, Saskatoon, Saskatchewan S7N 0S9 (CA).

(74) Agent: SINGLEHURST, John, C.; Finlayson & Singlehurst, 70 Gloucester Street, Ottawa, Ontario K2P 0A2 (CA).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: WHEAT PLANTS HAVING INCREASED RESISTANCE TO IMIDAZOLINONE HERBICIDES

Cross		Resistant (R)	Intermediate (I)	Susceptible (S)	Total F_1 screened R	I	S
Teal		0	0	334			
1A		59	0	0			
	Teal x 1A	0	5	0	0	10	0
	1A x Teal	0	5	0			
9A		66	0	0			
	Teal x 9A	0	5	0	0	10	0
	9A x Teal	0	5	0			
10A		66	0	0			
	Teal x 10A	0	5	0	0	11	0
	10A x Teal	0	6	0			
11A		53	0	0			
	Teal x 11A	0	7	0	0	15	0
	11A x Teal	0	8	0			
15A		48	0	0			
	Teal x 15A	7	0	0	14	0	0
	15A x Teal	7	0	0			
16A		66	0	0			
	Teal x 16A	0	7	0	0	14	0
	16A x Teal	0	7	0			

(57) Abstract: The present invention is directed to wheat plants having increased resistance to an imidazolinone herbicide. More particularly, the present invention includes wheat plants containing one or more IMI nucleic acids such as a Teal IMI cultivar. The nucleic acids are preferably located on or derived from different genomes. The present invention also includes seeds produced by these wheat plants and methods of controlling weeds in the vicinity of these wheat plants.

WO 03/014357 A1

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii)) for all designations*

— *of inventorship (Rule 4.17(iv)) for US only*

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WHEAT PLANTS HAVING INCREASED RESISTANCE TO IMIDAZOLINONE HERBICIDES

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the priority benefit of U.S. Provisional Application Serial No. 60/311,282 filed August 9, 2001.

FIELD OF THE INVENTION

The present invention relates in general to plants having an increased resistance to imidazolinone herbicides. More specifically, the present invention relates to wheat plants obtained by mutagenesis and cross-breeding and transformation that have an increased resistance to imidazolinone herbicides.

BACKGROUND OF THE INVENTION

Acetohydroxyacid synthase (AHAS; EC 4.1.3.18) is the first enzyme that catalyzes the biochemical synthesis of the branched chain amino acids valine, leucine and isoleucine (Singh B. K., 1999 Biosynthesis of valine, leucine and isoleucine in: Singh B. K. (Ed) Plant amino acids. Marcel Dekker Inc. New York, New York. Pg 227-247). AHAS is the site of action of four structurally diverse herbicide families including the sulfonylureas (LaRossa RA and Falco SC, 1984 Trends Biotechnol 2:158-161), the imidazolinones (Shaner et al., 1984 Plant Physiol 76:545-546), the triazolopyrimidines (Subramanian and Gerwick, 1989 Inhibition of acetolactate synthase by triazolopyrimidines in (ed) Whitaker JR, Sonnet PE Biocatalysis in agricultural biotechnology. ACS Symposium Series, American Chemical Society. Washington, D.C. Pg 277-288), and the pyrimidyloxybenzoates (Subramanian ét al., 1990 Plant Physiol 94: 239-244.). Imidazolinone and sulfonylurea herbicides are widely used in modern agriculture due to their effectiveness at very low application rates and relative non-toxicity in animals. By inhibiting AHAS activity, these families of herbicides prevent further growth and development of susceptible plants including many weed species. Several examples of commercially available imidazolinone herbicides are PURSUIT® (imazethapyr), SCEPTER® (imazaquin) and ARSENAL® (imazapyr). Examples of sulfonylurea herbicides are chlorsulfuron, metsulfuron methyl, sulfometuron methyl, chlorimuron

ethyl, thifensulfuron methyl, tribenuron methyl, bensulfuron methyl, nicosulfuron, ethametsulfuron methyl, rimsulfuron, triflusulfuron methyl, triasulfuron, primisulfuron methyl, cinosulfuron, amidosulfuron, fluzasulfuron, imazosulfuron, pyrazosulfuron ethyl and halosulfuron.

Due to their high effectiveness and low-toxicity, imidazolinone herbicides are favored for application by spraying over the top of a wide area of vegetation. The ability to spray an herbicide over the top of a wide range of vegetation decreases the costs associated with plantation establishment and maintenance and decreases the need for site preparation prior to use of such chemicals. Spraying over the top of a desired tolerant species also results in the ability to achieve maximum yield potential of the desired species due to the absence of competitive species. However, the ability to use such spray-over techniques is dependent upon the presence of imidazolinone resistant species of the desired vegetation in the spray over area.

Among the major agricultural crops, some leguminous species such as soybean are naturally resistant to imidazolinone herbicides due to their ability to rapidly metabolize the herbicide compounds (Shaner and Robinson, 1985 Weed Sci. 33:469-471). Other crops such as corn (Newhouse et al., 1992 Plant Physiol. 100:882-886) and rice (Barrette et al., 1989 Crop Safeners for Herbicides, Academic Press New York, pp. 195-220) are somewhat susceptible to imidazolinone herbicides. The differential sensitivity to the imidazolinone herbicides is dependent on the chemical nature of the particular herbicide and differential metabolism of the compound from a toxic to a non-toxic form in each plant (Shaner et al., 1984 Plant Physiol. 76:545-546; Brown et al., 1987 Pestic. Biochm. Physiol. 27:24-29). Other plant physiological differences such as absorption and translocation also play an important role in sensitivity (Shaner and Robinson, 1985 Weed Sci. 33:469-471).

Crop cultivars resistant to imidazolinones, sulfonylureas and triazolopyrimidines have been successfully produced using seed, microspore, pollen, and callus mutagenesis in *Zea mays*, *Arabidopsis thaliana*, *Brassica napus*, *Glycine max*, and *Nicotiana tabacum* (Sebastian et al., 1989 Crop Sci. 29:1403-1408; Swanson et al., 1989 Theor. Appl. Genet. 78:525-530; Newhouse et al., 1991 Theor. Appl. Genet. 83:65-70; Sathasivan et al., 1991 Plant Physiol. 97:1044-1050; Mourand et al., 1993 J. Heredity 84: 91-96). In all cases, a single, partially dominant nuclear

gene conferred resistance. Four imidazolinone resistant wheat plants were also previously isolated following seed mutagenesis of *Triticum aestivum* L. cv Fidel (Newhouse et al., 1992 Plant Physiol. 100:882-886). Inheritance studies confirmed that a single, partially dominant gene conferred resistance. Based on allelic studies, the authors concluded that the mutations in the four identified lines were located at the same locus. One of the Fidel cultivar resistance genes was designated FS-4 (Newhouse et al., 1992 Plant Physiol. 100:882-886).

Computer-based modeling of the three dimensional conformation of the AHAS-inhibitor complex predicts several amino acids in the proposed inhibitor binding pocket as sites where induced mutations would likely confer selective resistance to imidazolinones (Ott et al., 1996 J. Mol. Biol. 263:359-368) Wheat plants produced with some of these rationally designed mutations in the proposed binding sites of the AHAS enzyme have in fact exhibited specific resistance to a single class of herbicides (Ott et al., 1996 J. Mol. Biol. 263:359-368).

Plant resistance to imidazolinone herbicides has also been reported in a number of patents. U.S. Patent Nos. 4,761,373, 5,331,107, 5,304,732, 6,211,438, 6,211,439 and 6,222,100 generally describe the use of an altered AHAS gene to elicit herbicide resistance in plants, and specifically discloses certain imidazolinone resistant corn lines. U.S. Patent No. 5,013,659 discloses plants exhibiting herbicide resistance possessing mutations in at least one amino acid in one or more conserved regions. The mutations described therein encode either cross-resistance for imidazolinones and sulfonylureas or sulfonylurea-specific resistance, but imidazolinone-specific resistance is not described. Additionally, U.S. Patent No. 5,731,180 and U.S. Patent No. 5,767,361 discuss an isolated gene having a single amino acid substitution in a wild-type monocot AHAS amino acid sequence that results in imidazolinone-specific resistance.

To date, the prior art has not described imidazolinone resistant wheat plants containing more than one altered AHAS gene. Nor has the prior art described imidazolinone resistant wheat plants containing mutations on genomes other than the genome from which the FS-4 gene is derived. Therefore, what is needed in the art is the identification of imidazolinone resistance genes from additional genomes. What are also needed in the art are wheat plants having increased resistance to herbicides such as imidazolinone and containing more than one altered AHAS gene. Also

needed are methods for controlling weed growth in the vicinity of such wheat plants. These compositions and methods would allow for the use of spray over techniques when applying herbicides to areas containing wheat plants.

SUMMARY OF THE INVENTION

The present invention provides wheat plants comprising IMI nucleic acids, wherein the wheat plant has increased resistance to an imidazolinone herbicide as compared to a wild-type variety of the plant. The wheat plants can contain one, two, three or more IMI nucleic acids. In one embodiment, the wheat plant comprises multiple IMI nucleic acids located on different genomes. Preferably, the IMI nucleic acids encode proteins comprising a mutation in a conserved amino acid sequence selected from the group consisting of a Domain A, a Domain B, a Domain C, a Domain D and a Domain E. More preferably, the mutation is in a conserved Domain E or a conserved Domain C. Also provided are plant parts and plant seeds derived from the wheat plants described herein. In another embodiment, the wheat plant comprises an IMI nucleic acid that is not an Imi1 nucleic acid. The IMI nucleic acid can be an Imi2 or Imi3 nucleic acid, for example.

The IMI nucleic acids of the present invention can comprise a nucleotide sequence selected from the group consisting of: a polynucleotide of SEQ ID NO:1; a polynucleotide of SEQ ID NO:3; a polynucleotide sequence encoding a polypeptide of SEQ ID NO:2; a polynucleotide sequence encoding a polypeptide of SEQ ID NO:4, a polynucleotide comprising at least 60 consecutive nucleotides of any of the aforementioned polynucleotides; and a polynucleotide complementary to any of the aforementioned polynucleotides.

The plants of the present invention can be transgenic or non-transgenic. Examples of non-transgenic wheat plants having increased resistance to imidazolinone herbicides include a wheat plant having an ATCC Patent Deposit Designation Number PTA-3953 or PTA-3955; or a mutant, recombinant, or genetically engineered derivative of the plant with ATCC Patent Deposit Designation Number PTA-3953 or PTA-3955; or of any progeny of the plant with ATCC Patent Deposit Designation Number PTA-3953 or PTA-3955; or a plant that is a progeny of any of these plants.

In addition to the compositions of the present invention, several methods are provided. Described herein are methods of modifying a plant's tolerance to an imidazolinone herbicide comprising modifying the expression of an IMI nucleic acid in the plant. Also described are methods of producing a transgenic plant having increased tolerance to an imidazolinone herbicide comprising, transforming a plant cell with an expression vector comprising one or more IMI nucleic acids and generating the plant from the plant cell. The invention further includes a method of controlling weeds within the vicinity of a wheat plant, comprising applying an imidazolinone herbicide to the weeds and to the wheat plant, wherein the wheat plant has increased resistance to the imidazolinone herbicide as compared to a wild type variety of the wheat plant and wherein the plant comprises one or more IMI nucleic acids. In some preferred embodiments of these methods, the plants comprise multiple IMI nucleic acids that are located on different wheat genomes.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a table showing the results of single plant evaluation of imazamox resistance in parental and F_1 populations resulting from reciprocal crosses between resistant lines and CDC Teal. The numbers presented represent the number of plants scored into each phenotypic class. Parental lines are indicated in bold. The number of parental lines scored include those scored with the F_2 populations.

Figure 2 is a table showing the reaction to imazamox in F_2 and BCF_1 populations resulting from crosses between resistant lines and CDC Teal and Chi-square tests of single locus and two locus models (15A x Teal) for control of resistance. The symbols used in Figure 2 indicate the following: a - Chi-square P value (1 df) represents the probability that deviations from the tested ratio are due to chance alone. Chi-square P values greater than 0.05 indicate that observed values were not significantly different from expected values; b - Chi-square P value representing the probability that deviations between F_2 populations resulting from reciprocal crosses between CDC Teal and resistant lines are due to chance alone. Chi-square values greater than 0.05 indicate that reciprocal F_2 populations were homogeneous, and data from the two reciprocal populations was pooled; c - CDC Teal was used as the recurrent parent; d - Ratios tested were based on the results of the F_2 generation; and e – Chi-square P value (1 df) for BCF_1 ratio.

Figure 3 is a table showing the results of an evaluation of resistance to imazamox in $F_{2:3}$ families resulting from crosses between resistant lines and CDC Teal and Chi-square tests of single-locus and two locus models (15A x Teal) for control of resistance. The symbols used in Figure 3 indicate the following: a - Family segregation ratios tested were based on the results of the F_2 and BCF_1 populations; b - Chi-square P value (2 df) representing the probability that deviations from the tested ratio are due to chance alone. Chi-square P values greater than 0.05 indicate that observed values were not significantly different from expected values.

Figure 4 is a table showing the results of a single plant evaluation of imazamox resistance in F_2 populations resulting from inter-crosses between resistant lines. Chi-square ratios tested were based on the results of the F_2 and $F_{2:3}$ family results obtained from crosses between resistant lines and CDC Teal. The 15:1 ratio tested is for a two locus model and the 63:1 ratio tested is for a three locus model. The "a" symbol used in Figure 4 indicates the following: Chi-square P value (1 df) representing the probability that deviations from the tested ratio are due to chance alone. Chi-square P values greater than 0.05 indicate that observed values were not significantly different from expected values.

Figure 5 is a table showing the results of an evaluation of imazamox resistance in $F_{2:3}$ families resulting from segregating inter-crosses between resistant lines. The symbols used in Figure 5 indicate the following: a - Family segregation ratios tested were based on the results of the F_2 populations examined; b - Chi-square P value (2 df) representing the probability that deviations from the tested ratio are due to chance alone. Chi-square P values greater than 0.05 indicate that observed values were not significantly different from expected values.

Figure 6 is a table comparing the percent uninhibited in vitro AHAS activity in four wheat lines in the presence of increasing concentrations of the imidazolinone herbicide imazamox. Teal is a wild type line with no tolerance to imidazolinone herbicides while BW755 contains the FS4 mutant gene.

Figure 7 is a table comparing injury sustained by three wheat genotypes when treated with either a 10X or 30X rate of imazamox. The 1X rate is 20 g/ha. BW755 contains the FS4 mutant gene. 15A/11A is a bulk of selfed progeny from the cross of Teal11A and Teal15A. The population was not yet homozygous at all three non-allelic loci.

Figure 8 shows a DNA sequence alignment of partial *Als1* and *Imi1* wheat genes amplified from genomic DNA: CDC Teal (row 2; SEQ ID NO:15 and SEQ ID NO:16), BW755 (row 3; SEQ ID NO:17 and SEQ ID NO:18), TealIMI 10A (row 4; SEQ ID NO:19 and SEQ ID NO:20), TealIMI 11A (row 5; SEQ ID NO:21 and SEQ ID NO:22), and TealIMI 15A (row 6; SEQ ID NO:23 and SEQ ID NO:24). Partial sequences were aligned with a complete rice ALS gene sequence (row 1; SEQ ID NO:13 and SEQ ID NO:14) derived from Genbank (Accession no. ABO49822) and translated to protein sequences (presented on top of the DNA sequences). The five highly conserved amino acid domains known to house mutations that confer resistance to AHAS inhibitors are indicated in bold. Note the guanine to adenine substitutions in BW755, TealIMI 10A, and TealIMI 15A result in a serine to asparagine substitution (serine627 in rice) in the IPSGG domain (Domain E) of the *Als1* gene. Accordingly, the resistance genes present in the BW755, TealIMI 10A, and TealIMI 15A plants have been designated as part of the *Imi1* class. These Teal resistance genes are referred to herein as TealIMI1 10A and TealIMI1 15A.

Figure 9 shows a DNA sequence alignment of partial *Als2* and *Imi2* wheat genes amplified from genomic DNA: CDC Teal (row 2; SEQ ID NO:25 and SEQ ID NO:26), BW755 (row 3; SEQ ID NO:27 and SEQ ID NO:28), TealIMI 10A (row 4; SEQ ID NO:29 and SEQ ID NO:30), TealIMI 11A (row 5; SEQ ID NO:31 and SEQ ID NO:32) and TealIMI 15A (row 6; SEQ ID NO:33 and SEQ ID NO:34). Partial AHAS sequences were aligned with a complete rice AHAS sequence (row 1; SEQ ID NO:13 and SEQ ID NO:14) derived from GenBank (Accession no. AB049822) and translated into protein sequences (presented above the DNA sequences). The five highly conserved domains known to house mutations that confer resistance to AHAS inhibitors are indicated in bold. Note the guanine to adenine substitution in TealIMI 11A, resulting in a serine to asparagine substitution (serine627 in rice) in the IPSGG domain of the *Als2* gene. Accordingly, the resistance gene present in TealIMI 11A plant has been designated as part of the *Imi2* class of nucleic acids. This Teal resistance gene is referred to herein as TealIMI2 11A.

Figure 10 shows the partial DNA sequence of TealIMI1 15A (SEQ ID NO:1) and the deduced amino acid sequence of the same (SEQ ID NO:2).

Figure 11 shows the partial DNA sequence of TealIMI2 11A (SEQ ID NO:3) and the deduced amino acid sequence of the same (SEQ ID NO:4).

Figure 12 shows the wild type nucleic acid sequence of the Teal ALS1 ORF (SEQ ID NO:5), the Teal ALS2 ORF (SEQ ID NO:6) the Teal ALS3 ORF (SEQ ID NO:7).

Figure 13 is a schematic representation of the conserved amino acid sequences in the AHAS genes implicated in resistance to various AHAS inhibitors. The specific amino acid site responsible for resistance is indicated by an underline. (Modified from Devine, M. D. and Eberlein, C. V., 1997 Physiological, biochemical and molecular aspects of herbicide resistance based on altered target sites *in* Herbicide Activity: Toxicity, Biochemistry, and Molecular Biology, IOS Press Amsterdam, p. 159-185).

DETAILED DESCRIPTION

The present invention is directed to wheat plants, wheat plant parts and wheat plant cells having increased resistance to imidazolinone herbicides. The present invention also includes seeds produced by the wheat plants described herein and methods for controlling weeds in the vicinity of the wheat plants described herein. It is to be understood that as used in the specification and in the claims, "a" or "an" can mean one or more, depending upon the context in which it is used. Thus, for example, reference to "a cell" can mean that at least one cell can be utilized.

As used herein, the term "wheat plant" refers to a plant that is a member of the *Triticum* genus. The wheat plants of the present invention can be members of a *Triticum* genus including, but not limited to, *T. aestivum*, *T. turgidum*, *T. timopheevii*, *T. monococcum*, *T. zhukovskyi* and *T. urartu* and hybrids thereof. Examples of *T. aestivum* subspecies included within the present invention are *aestivum* (common wheat), *compactum* (club wheat), *macha* (macha wheat), *vavilovi* (vavilovi wheat), *spelta* and *sphaecrococcum* (shot wheat). Examples of *T. turgidum* subspecies included within the present invention are *turgidum*, *carthlicum*, *dicoccon*, *durum*, *paleocolchicum*, *polonicum*, *turanicum* and *dicoccoides*. Examples of *T. monococcum* subspecies included within the present invention are *monococcum* (einkorn) and *aegilopoides*. In one embodiment of the present invention, the wheat plant is a member of the *Triticum aestivum* species, and more particularly, the CDC Teal cultivar.

The term "wheat plant" is intended to encompass wheat plants at any stage of maturity or development as well as any tissues or organs (plant parts) taken or derived from any such plant unless otherwise clearly indicated by context. Plant parts include, but are not limited to, stems, roots, flowers, ovules, stamens, leaves, embryos, meristematic regions, callus tissue, anther cultures, gametophytes, sporophytes, pollen, microspores, protoplasts and the like. The present invention also includes seeds produced by the wheat plants of the present invention. In one embodiment, the seeds are true breeding for an increased resistance to an imidazolinone herbicide as compared to a wild type variety of the wheat plant seed.

The present invention describes a wheat plant comprising one or more IMI nucleic acids, wherein the wheat plant has increased resistance to an imidazolinone herbicide as compared to a wild-type variety of the plant. As used herein, the term "IMI nucleic acid" refers to a nucleic acid that is mutated from an AHAS nucleic acid in a wild type wheat plant that confers increased imidazolinone resistance to a plant in which it is transcribed. Wild type Teal AHAS nucleic acids are shown in SEQ ID NO:5 (Teal ALS1 ORF), SEQ ID NO:6 (Teal ALS2 ORF) and SEQ ID NO:7 (Teal ALS3 ORF). In one embodiment, the wheat plant comprises multiple IMI nucleic acids. As used when describing the IMI nucleic acids, the term "multiple" refers to IMI nucleic acids that have different nucleotide sequences and does not refer to a mere increase in number of the same IMI nucleic acid. For example, the IMI nucleic acids can be different due to the fact that they are derived from or located on different wheat genomes.

It is possible for the wheat plants of the present invention to have multiple IMI nucleic acids from different genomes since these plants can contain more than one genome. For example, a *Triticum aestivum* wheat plant contains three genomes sometimes referred to as the A, B and D genomes. Because AHAS is a required metabolic enzyme, it is assumed that each genome has at least one gene coding for the AHAS enzyme, commonly seen with other metabolic enzymes in hexaploid wheat that have been mapped. The AHAS nucleic acid on each genome can, and usually does, differ in its nucleotide sequence from an AHAS nucleic acid on another genome. One of skill in the art can determine the genome of origin of each AHAS nucleic acid through genetic crossing and/or either sequencing methods or exonuclease digestion methods known to those of skill in the art and as also described

in Example 2 below. For the purposes of this invention, IMI nucleic acids derived from one of the A, B or D genomes are distinguished and designated as Imi1, Imi2 or Imi3 nucleic acids.

It is not stated herein that any particular Imi nucleic acid class correlates with any particular A, B or D genome. For example, it is not stated herein that the Imi1 nucleic acids correlate to A genome nucleic acids, that Imi2 nucleic acids correlate to B genome nucleic acids, etc. The Imi1, Imi2 and Imi3 designations merely indicate that the IMI nucleic acids within each such class do not segregate independently, whereas two IMI nucleic acids from different classes do segregate independently and may therefore be derived from different wheat genomes. The Imi1 class of nucleic acids includes the FS-4 gene as described by Newhouse et al. (1992 Plant Physiol. 100:882-886) and the TealIMI1 15A gene described in more detail below. The Imi2 class of nucleic acids includes the TealIMI2 11A gene described below. Each Imi class can include members from different wheat species. Therefore, each Imi class includes IMI nucleic acids that differ in their nucleotide sequence but that are nevertheless designated as originating from, or being located on, the same wheat genome using inheritance studies as described in the Examples below and known to those of ordinary skill in the art.

Accordingly, the present invention includes a wheat plant comprising one or more IMI nucleic acids, wherein the wheat plant has increased resistance to an imidazolinone herbicide as compared to a wild-type variety of the plant and wherein the one or more IMI nucleic acids are selected from a group consisting of an Imi1, Imi2 and Imi3 nucleic acid. In one embodiment, the plant comprises an Imi1 nucleic and an Imi3 nucleic acid. In a preferred embodiment, the Imi1 nucleic acid comprises the polynucleotide sequence shown in SEQ ID NO:1. In another embodiment, the plant comprises an Imi2 nucleic acid. In a preferred embodiment, the Imi2 nucleic acid comprises the polynucleotide sequence shown in SEQ ID NO:3.

As used herein with regard to nucleic acids, the term "from" refers to a nucleic acid "located on" or "derived from" a particular genome. The term "located on" refers to a nucleic acid contained within that particular genome. As also used herein with regard to a genome, the term "derived from" refers to a nucleic acid that has been removed or isolated from that genome. The term "isolated" is defined in more detail below.

In another embodiment, the wheat plant comprises an IMI nucleic acid, wherein the nucleic acid is a non-Imi1 nucleic acid. The term "non-Imi1", refers to an IMI nucleic acid that is not a member of the Imi1 class as described above. Examples of nucleic acids from the Imi1 class are shown in rows 3, 4 and 5 of Figure 8. One example of non-Imi1 nucleic acid is shown in row 5 of Figure 8. Accordingly, in a preferred embodiment, the wheat plant comprises an IMI nucleic acid comprising the polynucleotide sequence encoding the polypeptide of SEQ ID NO:4. The polynucleotide sequence can comprise the sequence shown in SEQ ID NO:3.

The present invention includes wheat plants comprising one, two, three or more IMI nucleic acids, wherein the wheat plant has increased resistance to an imidazolinone herbicide as compared to a wild-type variety of the plant. The IMI nucleic acids can comprise a nucleotide sequence selected from the group consisting of a polynucleotide of SEQ ID NO:1; a polynucleotide of SEQ ID NO:3; a polynucleotide sequence encoding a polypeptide of SEQ ID NO:2; a polynucleotide sequence encoding a polypeptide of SEQ ID NO:4, a polynucleotide comprising at least 60 consecutive nucleotides of any of the aforementioned polynucleotides; and a polynucleotide complementary to any of the aforementioned polynucleotides.

The imidazolinone herbicide can be selected from, but is not limited to, PURSUIT® (imazethapyr), CADRE® (imazapic), RAPTOR® (imazamox), SCEPTER® (imazaquin), ASSERT® (imazethabenz), ARSENAL® (imazapyr), a derivative of any of the aforementioned herbicides, or a mixture of two or more of the aforementioned herbicides, for example, imazapyr/imazamox (ODYSSEY®). More specifically, the imidazolinone herbicide can be selected from, but is not limited to, 2-(4-isopropyl-4-methyl-5-oxo-2-imidiazolin-2-y1)-nicotinic acid, 2-(4-isopropyl)-4-methyl-5-oxo-2-imidazolin-2-y1)-3-quinolinecarboxylic acid, 5-ethyl-2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-y1)-nicotinic acid, 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-5-(methoxymethyl)-nicotinic acid, 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-5-methylnicotinic acid, and a mixture of methyl 6-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-m-toluate and methyl 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-p-toluate. The use of 5-ethyl-2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-nicotinic acid and 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-5-(methoxymethyl)-nicotinic acid is preferred. The use of 2-(4-isopropyl-4-

methyl-5-oxo-2-imidazolin-2-yl)-5-(methoxymethyl)-nicotinic acid is particularly preferred.

In one embodiment, the wheat plant comprises two IMI nucleic acids, wherein the nucleic acids are derived from or located on different wheat genomes. Preferably, one of the two nucleic acids is an Imi1 nucleic acid, and more preferably comprises the polynucleotide sequence of SEQ ID NO:1. In another embodiment, the wheat plant comprises one IMI nucleic acid, wherein the nucleic acid comprises the polynucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3. In yet another embodiment, the wheat plant comprises three or more IMI nucleic acids wherein each nucleic acid is from a different genome. Preferably, at least one of the three IMI nucleic acids comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:1 and SEQ ID NO:3.

In a preferred embodiment of the present invention, the one or more IMI nucleic acids contained within the plant encode an amino acid sequence comprising a mutation in a domain that is conserved among several AHAS proteins. These conserved domains are referred to herein as Domain A, Domain B, Domain C, Domain D and Domain E. Figure 13 shows the general location of each domain in an AHAS protein. As used herein, Domain A contains the amino acid sequence AITGQVPRRMIGT (SEQ ID NO:8); Domain B contains the amino acid sequence QWED (SEQ ID NO:9); Domain C contains the amino acid sequence VFAYPGGASMEIHQALTRS (SEQ ID NO:10); Domain D contains the amino acid sequence AFQETP (SEQ ID NO:11); Domain E contains the amino acid sequence IPSGG (SEQ ID NO:12). The present invention also contemplates that there may be slight variations in the conserved domains, for example, in cockleberry plants, the serine residue in Domain E is replaced by an alanine residue.

Accordingly, the present invention includes a wheat plant comprising an IMI nucleic acid that encodes an amino acid sequence having a mutation in a conserved domain selected from the group consisting of a Domain A, a Domain B, a Domain C, a Domain D and a Domain E. In one embodiment, the wheat plant comprises an IMI nucleic acid that encodes an amino acid sequence having a mutation in a Domain E. In further preferred embodiments, the mutations in the conserved domains occur at the locations indicated by the following underlining: AITGQVPRRMIGT (SEQ ID NO:8); QWED (SEQ ID NO:9);

VFAYPGGASMEIHQALTRS (SEQ ID NO:10); AFQETP (SEQ ID NO:11) and IPSGG (SEQ ID NO:12). One preferred substitution is asparagine for serine in Domain E (SEQ ID NO:12).

The wheat plants described herein can be either transgenic wheat plants or non-transgenic wheat plants. As used herein, the term "transgenic" refers to any plant, plant cell, callus, plant tissue or plant part, that contains all or part of at least one recombinant polynucleotide. In many cases, all or part of the recombinant polynucleotide is stably integrated into a chromosome or stable extra-chromosomal element, so that it is passed on to successive generations. For the purposes of the invention, the term "recombinant polynucleotide" refers to a polynucleotide that has been altered, rearranged or modified by genetic engineering. Examples include any cloned polynucleotide, or polynucleotides, that are linked or joined to heterologous sequences. The term "recombinant" does not refer to alterations of polynucleotides that result from naturally occurring events, such as spontaneous mutations, or from non-spontaneous mutagenesis followed by selective breeding. Plants containing mutations arising due to non-spontaneous mutagenesis and selective breeding are referred to herein as non-transgenic plants and are included in the present invention. In embodiments wherein the wheat plant is transgenic and comprises multiple IMI nucleic acids, the nucleic acids can be derived from different genomes or the same genome. Alternatively, in embodiments wherein the wheat plant is non-transgenic and comprises multiple IMI nucleic acids, the nucleic acids are located on different genomes.

An example of a non-transgenic wheat plant cultivar comprising one IMI nucleic acid is the plant cultivar deposited with the ATCC under Patent Deposit Designation Number PTA-3953 and designated herein as the TealIMI 11A wheat cultivar. The TealIMI 11A wheat cultivar contains an Imi2 nucleic acid. The partial nucleotide and deduced amino acid sequences corresponding to the TealIMI2 11A gene are shown in SEQ ID NO:3 and SEQ ID NO:4, respectively. The only portion of the sequences not included in SEQ ID NO:3 and SEQ ID NO:4 are those sequences encoding and corresponding to a signal sequence that is cleaved from the mature TealIMI2 11A protein. Accordingly, SEQ ID NO:4 represents the full deduced sequence of the mature TealIMI2 11A protein.

An example of a wheat plant cultivar comprising two IMI nucleic acids on different genomes is the plant cultivar deposited with the ATCC under Patent Deposit Designation Number PTA-3955 and designated herein as the TealIMI 15A wheat cultivar. The TealIMI 15A wheat cultivar contains Imi1 and Imi3 nucleic acids. The Imi1 nucleic acid comprises a mutation that results in a serine to asparagine change in the IMI protein encoded thereby. The mutated AHAS genes are designated herein as TealIMI1 15A and TealIMI3 15A. The partial nucleotide and deduced amino acid sequences corresponding to the TealIMI1 15A gene are shown in SEQ ID NO:1 and SEQ ID NO:2, respectively. The only portion of the sequences not included in SEQ ID NO:1 and SEQ ID NO:2 are those sequences encoding and corresponding to approximately 100-150 base pairs at the 5' end and approximately 5 base pairs at the 3' end of the coding region.

Separate deposits of 2500 seeds of the TealIMI 11A and TealIMI 15A wheat cultivars were made with the American Type Culture Collection, Manassas, Virginia on January 3, 2002. These deposits were made in accordance with the terms and provisions of the Budapest Treaty relating to the deposit of microorganisms. The deposits were made for a term of at least thirty years and at least five years after the most recent request for the furnishing of a sample of the deposit is received by the ATCC. The deposited seeds were accorded Patent Deposit Designation Numbers PTA-3953 (TealIMI 11A) and PTA-3955 (TealIMI 15A).

The present invention includes the wheat plant having a Patent Deposit Designation Number PTA-3953 or PTA-3955; a mutant, recombinant, or genetically engineered derivative of the plant with Patent Deposit Designation Number PTA-3953 or PTA-3955; any progeny of the plant with Patent Deposit Designation Number PTA-3953 or PTA-3955; and a plant that is the progeny of any of these plants. In a preferred embodiment, the wheat plant of the present invention additionally has the herbicide resistance characteristics of the plant with Patent Deposit Designation Number PTA-3953 or PTA-3955.

Also included in the present invention are hybrids of the TealIMI 11A and TealIMI 15A wheat cultivars described herein. Example 5 demonstrates TealIMI11A/TealIMI15A hybrids having increased resistance to an imidazolinone herbicide. The present invention further includes hybrids of the TealIMI 11A or TealIMI 15A wheat cultivars and another wheat cultivar. The other wheat cultivar

includes, but is not limited to, *T. aestivum* L. cv Fidel and any wheat cultivar harboring a mutant gene FS-1, FS-2, FS-3 or FS-4. (See U.S. Patent No. 6,339,184 and U.S. Patent Application No. 08/474,832). In a preferred embodiment, the wheat plant is a hybrid between a TealIMI 11A cultivar and a Fidel FS-4 cultivar. The TealIMI 11A/FS-4 hybrids comprise an Imi1 nucleic acid and an Imi2 nucleic acid. A hybrid of TealIMI 11A and a Fidel cultivar harboring the FS-4 gene is included in the present invention and was deposited with the American Type Culture Collection, Manassas, Virginia on January 3, 2002. This deposit was made in accordance with the terms and provisions of the Budapest Treaty relating to the deposit of microorganisms. The deposit was made for a term of at least thirty years and at least five years after the most recent request for the furnishing of a sample of the deposit is received by the ATCC. The deposited seeds were accorded Patent Deposit Designation Number PTA-3954.

The terms "cultivar" and "variety" refer to a group of plants within a species defined by the sharing of a common set of characteristics or traits accepted by those skilled in the art as sufficient to distinguish one cultivar or variety from another cultivar or variety. There is no implication in either term that all plants of any given cultivar or variety will be genetically identical at either the whole gene or molecular level or that any given plant will be homozygous at all loci. A cultivar or variety is considered "true breeding" for a particular trait if, when the true-breeding cultivar or variety is self-pollinated, all of the progeny contain the trait. In the present invention, the trait arises from a mutation in an AHAS gene of the wheat plant or seed.

It is to be understood that the wheat plant of the present invention can comprise a wild type or non-mutated AHAS gene in addition to an IMI gene. As described in Example 4, it is contemplated that wheat cultivar TealIMI 11A contains a mutation in only one of multiple AHAS isoenzymes and that wheat cultivar TealIMI 15A contains a mutation in only two of multiple AHAS isoenzymes. Therefore, the present invention includes a wheat plant comprising one or more IMI nucleic acids in addition to one or more wild type or non-mutated AHAS nucleic acids.

In addition to wheat plants, the present invention encompasses isolated IMI proteins and nucleic acids. The nucleic acids comprise a polynucleotide selected from the group consisting of a polynucleotide of SEQ ID NO:1; a polynucleotide of SEQ ID NO:3; a polynucleotide sequence encoding a polypeptide of SEQ ID NO:2; a

polynucleotide sequence encoding a polypeptide of SEQ ID NO:4, a polynucleotide comprising at least 60 consecutive nucleotides of any of the aforementioned polynucleotides; and a polynucleotide complementary to any of the aforementioned polynucleotides. In a preferred embodiment, the IMI nucleic acid comprises a polynucleotide sequence encoding a polypeptide of SEQ ID NO:2 or SEQ ID NO:4. In a further preferred embodiment, the IMI nucleic acid comprises a polynucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3.

The term "AHAS protein" refers to an acetohydroxyacid synthase protein and the term "IMI protein" refers to any AHAS protein that is mutated from a wild type AHAS protein and that confers increased imidazolinone resistance to a plant, plant cell, plant part, plant seed or plant tissue when it is expressed therein. In a preferred embodiment, the IMI protein comprises a polypeptide of SEQ ID NO:2 or SEQ ID NO:4. As also used herein, the terms "nucleic acid" and "polynucleotide" refer to RNA or DNA that is linear or branched, single or double stranded, or a hybrid thereof. The term also encompasses RNA/DNA hybrids. These terms also encompass untranslated sequence located at both the 3' and 5' ends of the coding region of the gene: at least about 1000 nucleotides of sequence upstream from the 5' end of the coding region and at least about 200 nucleotides of sequence downstream from the 3' end of the coding region of the gene. Less common bases, such as inosine, 5-methylcytosine, 6-methyladenine, hypoxanthine and others can also be used for antisense, dsRNA and ribozyme pairing. For example, polynucleotides that contain C-5 propyne analogues of uridine and cytidine have been shown to bind RNA with high affinity and to be potent antisense inhibitors of gene expression. Other modifications, such as modification to the phosphodiester backbone, or the 2'-hydroxy in the ribose sugar group of the RNA can also be made. The antisense polynucleotides and ribozymes can consist entirely of ribonucleotides, or can contain mixed ribonucleotides and deoxyribonucleotides. The polynucleotides of the invention may be produced by any means, including genomic preparations, cDNA preparations, *in vitro* synthesis, RT-PCR and *in vitro* or *in vivo* transcription.

An "isolated" nucleic acid molecule is one that is substantially separated from other nucleic acid molecules, which are present in the natural source of the nucleic acid (i.e., sequences encoding other polypeptides). Preferably, an "isolated" nucleic acid is free of some of the sequences that naturally flank the

nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in its naturally occurring replicon. For example, a cloned nucleic acid is considered isolated. In various embodiments, the isolated IMI nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (e.g., a *Triticum aestivum* cell). A nucleic acid is also considered isolated if it has been altered by human intervention, or placed in a locus or location that is not its natural site, or if it is introduced into a cell by agroinfection or biolistics. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be free from some of the other cellular material with which it is naturally associated, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized.

Specifically excluded from the definition of "isolated nucleic acids" are: naturally–occurring chromosomes (such as chromosome spreads), artificial chromosome libraries, genomic libraries, and cDNA libraries that exist either as an *in vitro* nucleic acid preparation or as a transfected/transformed host cell preparation, wherein the host cells are either an *in vitro* heterogeneous preparation or plated as a heterogeneous population of single colonies. Also specifically excluded are the above libraries wherein a specified nucleic acid makes up less than 5% of the number of nucleic acid inserts in the vector molecules. Further specifically excluded are whole cell genomic DNA or whole cell RNA preparations (including whole cell preparations that are mechanically sheared or enzymatically digested). Even further specifically excluded are the whole cell preparations found as either an *in vitro* preparation or as a heterogeneous mixture separated by electrophoresis wherein the nucleic acid of the invention has not further been separated from the heterologous nucleic acids in the electrophoresis medium (e.g., further separating by excising a single band from a heterogeneous band population in an agarose gel or nylon blot).

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule containing a nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3 or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, a *T. aestivum* IMI cDNA can be isolated from a *T. aestivum* library using all or a portion of the sequence of SEQ ID NO:1 or SEQ ID NO:3. Moreover, a nucleic acid molecule encompassing all or a

portion of SEQ ID NO:1 or SEQ ID NO:3 can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this sequence. For example, mRNA can be isolated from plant cells (e.g., by the guanidinium-thiocyanate extraction procedure of Chirgwin et al., 1979 Biochemistry 18:5294-5299) and cDNA can be prepared using reverse transcriptase (e.g., Moloney MLV reverse transcriptase, available from Gibco/BRL, Bethesda, MD; or AMV reverse transcriptase, available from Seikagaku America, Inc., St. Petersburg, FL). Synthetic oligonucleotide primers for polymerase chain reaction amplification can be designed based upon the nucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:3. A nucleic acid molecule of the invention can be amplified using cDNA or, alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid molecule so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to an IMI nucleotide sequence can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

The IMI nucleic acids of the present invention can comprise sequences encoding an IMI protein (i.e., "coding regions"), as well as 5' untranslated sequences and 3' untranslated sequences. Alternatively, the nucleic acid molecules of the present invention can comprise only the coding regions of an IMI gene, or can contain whole genomic fragments isolated from genomic DNA. A coding region of these sequences is indicated as an "ORF position". Moreover, the nucleic acid molecule of the invention can comprise a portion of a coding region of an IMI gene, for example, a fragment that can be used as a probe or primer. The nucleotide sequences determined from the cloning of the IMI genes from *T. aestivum* allow for the generation of probes and primers designed for use in identifying and/or cloning IMI homologs in other cell types and organisms, as well as IMI homologs from other wheat plants and related species. The portion of the coding region can also encode a biologically active fragment of an IMI protein.

As used herein, the term "biologically active portion of" an IMI protein is intended to include a portion, e.g., a domain/motif, of an IMI protein that, when produced in a plant increases the plant's resistance to an imidazolinone herbicide as compared to a wild-type variety of the plant. Methods for quantitating increased resistance to imidazolinone herbicides are provided in the Examples

provided below. Biologically active portions of an IMI protein include peptides encoded by polynucleotide sequences comprising SEQ ID NO:1 or SEQ ID NO:3 which include fewer amino acids than a full length IMI protein and impart increased resistance to an imidazolinone herbicide upon expression in a plant. Typically, biologically active portions (e.g., peptides which are, for example, 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) comprise a domain or motif with at least one activity of an IMI protein. Moreover, other biologically active portions in which other regions of the polypeptide are deleted, can be prepared by recombinant techniques and evaluated for one or more of the activities described herein. Preferably, the biologically active portions of an IMI protein include one or more conserved domains selected from the group consisting of a Domain A, a Domain B, a Domain C, a Domain D and a Domain E, wherein the conserved domain contains a mutation.

The invention also provides IMI chimeric or fusion polypeptides. As used herein, an IMI "chimeric polypeptide" or "fusion polypeptide" comprises an IMI polypeptide operatively linked to a non-IMI polypeptide. A "non-IMI polypeptide" refers to a polypeptide having an amino acid sequence that is not substantially identical to an IMI polypeptide, e.g., a polypeptide that is not an IMI isoenzyme, which peptide performs a different function than an IMI polypeptide. Within the fusion polypeptide, the term "operatively linked" is intended to indicate that the IMI polypeptide and the non-IMI polypeptide are fused to each other so that both sequences fulfill the proposed function attributed to the sequence used. The non-IMI polypeptide can be fused to the N-terminus or C-terminus of the IMI polypeptide. For example, in one embodiment, the fusion polypeptide is a GST-IMI fusion polypeptide in which the IMI sequence is fused to the C-terminus of the GST sequence. Such fusion polypeptides can facilitate the purification of recombinant IMI polypeptides. In another embodiment, the fusion polypeptide is an IMI polypeptide containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells), expression and/or secretion of an IMI polypeptide can be increased through use of a heterologous signal sequence.

An isolated nucleic acid molecule encoding an IMI polypeptide having sequence identity to a polypeptide encoded by a polynucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3 can be created by introducing one or more nucleotide

substitutions, additions or deletions into a nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3 such that one or more amino acid substitutions, additions or deletions are introduced into the encoded polypeptide. Mutations can be introduced into a sequence of SEQ ID NO:1 or SEQ ID NO:3 by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues.

A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in an IMI polypeptide is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an IMI coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for an IMI activity described herein to identify mutants that retain IMI activity. Following mutagenesis of the sequence of SEQ ID NO:1 or SEQ ID NO:3, the encoded polypeptide can be expressed recombinantly and the activity of the polypeptide can be determined by analyzing the imidazolinone resistance of a plant expressing the polypeptide as described in the Examples below.

To determine the percent sequence identity of two amino acid sequences (e.g., SEQ ID NO:2 or SEQ ID NO:4 and a mutant form thereof), the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of one polypeptide for optimal alignment with the other polypeptide). The amino acid residues at corresponding amino acid positions are then compared. When a position in one sequence (e.g., SEQ ID NO:2 or SEQ ID NO:4) is occupied by the same amino acid residue as the corresponding position in the other sequence (e.g., a mutant form of SEQ ID NO:2 or SEQ ID NO:4), then the molecules are

identical at that position. The same type of comparison can be made between two nucleic acid sequences. The percent sequence identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., percent sequence identity = numbers of identical positions/total numbers of positions x 100). For the purposes of the invention, the percent sequence identity between two nucleic acid or polypeptide sequences is determined using the Vector NTI 6.0 (PC) software package (InforMax, 7600 Wisconsin Ave., Bethesda, MD 20814). A gap opening penalty of 15 and a gap extension penalty of 6.66 are used for determining the percent identity of two nucleic acids. A gap opening penalty of 10 and a gap extension penalty of 0.1 are used for determining the percent identity of two polypeptides. All other parameters are set at the default settings. It is to be understood that for the purposes of determining sequence identity, when comparing a DNA sequence to an RNA sequence, a thymidine nucleotide is equivalent to a uracil nucleotide. Preferably, the isolated IMI polypeptides included in the present invention are at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-75%, 75-80%, 80-85%, 85-90% or 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more identical to an entire amino acid sequence encoded by a polynucleotide sequence shown in SEQ ID NO:1 or SEQ ID NO:3. In another embodiment, the isolated IMI polypeptides included in the present invention are at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-75%, 75-80%, 80-85%, 85-90% or 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more identical to an entire amino acid sequence shown in SEQ ID NO:2 or SEQ ID NO:4.

Additionally, optimized IMI nucleic acids can be created. Preferably, an optimized IMI nucleic acid encodes an IMI polypeptide that modulates a plant's tolerance to imidazolinone herbicides, and more preferably increases a plant's tolerance to an imidazolinone herbicide upon its over-expression in the plant. As used herein, "optimized" refers to a nucleic acid that is genetically engineered to increase its expression in a given plant or animal. To provide plant optimized IMI nucleic acids, the DNA sequence of the gene can be modified to 1) comprise codons preferred by highly expressed plant genes; 2) comprise an A+T content in nucleotide base composition to that substantially found in plants; 3) form a plant initiation sequence, 4) eliminate sequences that cause destabilization, inappropriate

polyadenylation, degradation and termination of RNA, or that form secondary structure hairpins or RNA splice sites. Increased expression of IMI nucleic acids in plants can be achieved by utilizing the distribution frequency of codon usage in plants in general or a particular plant. Methods for optimizing nucleic acid expression in plants can be found in EPA 0359472; EPA 0385962; PCT Application No. WO 91/16432; U.S. Patent No. 5,380,831; U.S. Patent No. 5,436,391; Perlack *et al.*, 1991 Proc. Natl. Acad. Sci. USA 88:3324-3328; and Murray *et al.*, 1989 Nucleic Acids Res. 17:477-498.

As used herein, "frequency of preferred codon usage" refers to the preference exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. To determine the frequency of usage of a particular codon in a gene, the number of occurrences of that codon in the gene is divided by the total number of occurrences of all codons specifying the same amino acid in the gene. Similarly, the frequency of preferred codon usage exhibited by a host cell can be calculated by averaging frequency of preferred codon usage in a large number of genes expressed by the host cell. It is preferable that this analysis be limited to genes that are highly expressed by the host cell. The percent deviation of the frequency of preferred codon usage for a synthetic gene from that employed by a host cell is calculated first by determining the percent deviation of the frequency of usage of a single codon from that of the host cell followed by obtaining the average deviation over all codons. As defined herein, this calculation includes unique codons (i.e., ATG and TGG). In general terms, the overall average deviation of the codon usage of an optimized gene from that of a host cell is calculated using the equation $1A = n = 1\ Z\ X_n - Y_n\ X_n$ times 100 Z where X_n = frequency of usage for codon n in the host cell; Y_n = frequency of usage for codon n in the synthetic gene, n represents an individual codon that specifies an amino acid and the total number of codons is Z. The overall deviation of the frequency of codon usage, A, for all amino acids should preferably be less than about 25%, and more preferably less than about 10%.

Hence, an IMI nucleic acid can be optimized such that its distribution frequency of codon usage deviates, preferably, no more than 25% from that of highly expressed plant genes and, more preferably, no more than about 10%. In addition, consideration is given to the percentage G+C content of the degenerate third base (monocotyledons appear to favor G+C in this position, whereas dicotyledons do not).

It is also recognized that the XCG (where X is A, T, C, or G) nucleotide is the least preferred codon in dicots whereas the XTA codon is avoided in both monocots and dicots. Optimized IMI nucleic acids of this invention also preferably have CG and TA doublet avoidance indices closely approximating those of the chosen host plant (i.e., *Triticum aestivum*). More preferably these indices deviate from that of the host by no more than about 10-15%.

In addition to the nucleic acid molecules encoding the IMI polypeptides described above, another aspect of the invention pertains to isolated nucleic acid molecules that are antisense thereto. Antisense polynucleotides are thought to inhibit gene expression of a target polynucleotide by specifically binding the target polynucleotide and interfering with transcription, splicing, transport, translation and/or stability of the target polynucleotide. Methods are described in the prior art for targeting the antisense polynucleotide to the chromosomal DNA, to a primary RNA transcript or to a processed mRNA. Preferably, the target regions include splice sites, translation initiation codons, translation termination codons, and other sequences within the open reading frame.

The term "antisense", for the purposes of the invention, refers to a nucleic acid comprising a polynucleotide that is sufficiently complementary to all or a portion of a gene, primary transcript or processed mRNA, so as to interfere with expression of the endogenous gene. "Complementary" polynucleotides are those that are capable of base pairing according to the standard Watson-Crick complementarity rules. Specifically, purines will base pair with pyrimidines to form a combination of guanine paired with cytosine (G:C) and adenine paired with either thymine (A:T) in the case of DNA, or adenine paired with uracil (A:U) in the case of RNA. It is understood that two polynucleotides may hybridize to each other even if they are not completely complementary to each other, provided that each has at least one region that is substantially complementary to the other. The term "antisense nucleic acid" includes single stranded RNA as well as double-stranded DNA expression cassettes that can be transcribed to produce an antisense RNA. "Active" antisense nucleic acids are antisense RNA molecules that are capable of selectively hybridizing with a primary transcript or mRNA encoding a polypeptide having at least 80% sequence identity with the polypeptide encoded by the polynucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3.

In addition to the IMI nucleic acids and polypeptides described above, the present invention encompasses these nucleic acids and polypeptides attached to a moiety. These moieties include, but are not limited to, detection moieties, hybridization moieties, purification moieties, delivery moieties, reaction moieties, binding moieties, and the like. A typical group of nucleic acids having moieties attached are probes and primers. Probes and primers typically comprise a substantially isolated oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 40, 50 or 75 consecutive nucleotides of a sense strand of the sequence set forth in SEQ ID NO:1 or SEQ ID NO:3, an anti-sense sequence of the sequence set forth in SEQ ID NO:1 or SEQ ID NO:3, or naturally occurring mutants thereof. Primers based on a nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3 can be used in PCR reactions to clone IMI homologs. Probes based on the IMI nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous polypeptides. In preferred embodiments, the probe further comprises a label group attached thereto, e.g. the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a genomic marker test kit for identifying cells which express an IMI polypeptide, such as by measuring a level of an IMI-encoding nucleic acid, in a sample of cells, e.g., detecting IMI mRNA levels or determining whether a genomic IMI gene has been mutated or deleted.

The invention further provides an isolated recombinant expression vector comprising an IMI nucleic acid as described above, wherein expression of the vector in a host cell results in increased resistance to an imidazolinone herbicide as compared to a wild type variety of the host cell. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon

introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operably linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/ translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990) or see: Gruber and Crosby, in: Methods in Plant Molecular Biology and Biotechnology, eds. Glick and Thompson, Chapter 7, 89-108, CRC Press: Boca Raton, Florida, including the references therein. Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cells and those that direct expression of the nucleotide sequence only in certain host cells or under certain conditions. It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of polypeptide desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce polypeptides or peptides, including fusion polypeptides or peptides, encoded

by nucleic acids as described herein (e.g., IMI polypeptides, fusion polypeptides, etc.).

In a preferred embodiment of the present invention, the IMI polypeptides are expressed in plants and plants cells such as unicellular plant cells (such as algae) (see Falciatore *et al.*, 1999 Marine Biotechnology 1(3):239-251 and references therein) and plant cells from higher plants (e.g., the spermatophytes, such as crop plants). An IMI polynucleotide may be "introduced" into a plant cell by any means, including transfection, transformation or transduction, electroporation, particle bombardment, biolistics, agroinfection and the like. One transformation method known to those of skill in the art is the dipping of a flowering plant into an Agrobacteria solution, wherein the Agrobacteria contains the IMI nucleic acid, followed by breeding of the transformed gametes.

Other suitable methods for transforming or transfecting host cells including plant cells can be found in Sambrook, et al. (Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) and other laboratory manuals such as Methods in Molecular Biology, 1995, Vol. 44, Agrobacterium protocols, ed: Gartland and Davey, Humana Press, Totowa, New Jersey. As increased resistance to imidazolinone herbicides is a general trait wished to be inherited into a wide variety of plants like maize, wheat, rye, oat, triticale, rice, barley, soybean, peanut, cotton, rapeseed and canola, manihot, pepper, sunflower and tagetes, solanaceous plants like potato, tobacco, eggplant, and tomato, Vicia species, pea, alfalfa, bushy plants (coffee, cacao, tea), Salix species, trees (oil palm, coconut), perennial grasses and forage crops, these crop plants are also preferred target plants for a genetic engineering as one further embodiment of the present invention. Forage crops include, but are not limited to, Wheatgrass, Canarygrass, Bromegrass, Wildrye Grass, Bluegrass, Orchardgrass, Alfalfa, Salfoin, Birdsfoot Trefoil, Alsike Clover, Red Clover and Sweet Clover.

In one embodiment of the present invention, transfection of an IMI polynucleotide into a plant is achieved by Agrobacterium mediated gene transfer. Agrobacterium mediated plant transformation can be performed using for example the GV3101(pMP90) (Koncz and Schell, 1986 Mol. Gen. Genet. 204:383-396) or LBA4404 (Clontech) *Agrobacterium tumefaciens* strain. Transformation can be

performed by standard transformation and regeneration techniques (Deblaere *et al.*, 1994 Nucl. Acids. Res. 13:4777-4788; Gelvin, Stanton B. and Schilperoort, Robert A, Plant Molecular Biology Manual, 2nd Ed. - Dordrecht : Kluwer Academic Publ., 1995. - in Sect., Ringbuc Zentrale Signatur: BT11-P ISBN 0-7923-2731-4; Glick, Bernard R. and Thompson, John E., Methods in Plant Molecular Biology and Biotechnology, Boca Raton : CRC Press, 1993 360 S., ISBN 0-8493-5164-2). For example, rapeseed can be transformed via cotyledon or hypocotyl transformation (Moloney *et al.*, 1989 Plant cell Report 8:238-242; De Block *et al.*, 1989 Plant Physiol. 91:694-701). Use of antibiotica for Agrobacterium and plant selection depends on the binary vector and the Agrobacterium strain used for transformation. Rapeseed selection is normally performed using kanamycin as selectable plant marker. Agrobacterium mediated gene transfer to flax can be performed using, for example, a technique described by Mlynarova *et al.*, 1994 Plant Cell Report 13:282-285. Additionally, transformation of soybean can be performed using, for example, a technique described in European Patent No. 0424 047, U.S. Patent No. 5,322,783, European Patent No. 0397 687, U.S. Patent No. 5,376,543 or U.S. Patent No. 5,169,770. Transformation of maize can be achieved by particle bombardment, polyethylene glycol mediated DNA uptake or via the silicon carbide fiber technique. (See, for example, Freeling and Walbot "The maize handbook" Springer Verlag: New York (1993) ISBN 3-540-97826-7). A specific example of maize transformation is found in U.S. Patent No. 5,990,387 and a specific example of wheat transformation can be found in PCT Application No. WO 93/07256.

According to the present invention, the introduced IMI polynucleotide may be maintained in the plant cell stably if it is incorporated into a non-chromosomal autonomous replicon or integrated into the plant chromosomes. Alternatively, the introduced IMI polynucleotide may be present on an extra-chromosomal non-replicating vector and be transiently expressed or transiently active. In one embodiment, a homologous recombinant microorganism can be created wherein the IMI polynucleotide is integrated into a chromosome, a vector is prepared which contains at least a portion of an AHAS gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the endogenous AHAS gene and to create an IMI gene. To create a point mutation via homologous recombination, DNA-RNA hybrids can be used in a technique known as

chimeraplasty (Cole-Strauss *et al.*, 1999 Nucleic Acids Research 27(5):1323-1330 and Kmiec, 1999 Gene therapy American Scientist 87(3):240-247). Other homologous recombination procedures in *Triticum* species are also well known in the art and are contemplated for use herein.

In the homologous recombination vector, the IMI gene can be flanked at its 5' and 3' ends by an additional nucleic acid molecule of the AHAS gene to allow for homologous recombination to occur between the exogenous IMI gene carried by the vector and an endogenous AHAS gene, in a microorganism or plant. The additional flanking AHAS nucleic acid molecule is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several hundreds of base pairs up to kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (see e.g., Thomas, K. R., and Capecchi, M. R., 1987 Cell 51:503 for a description of homologous recombination vectors or Strepp *et al.*, 1998 PNAS, 95(8):4368-4373 for cDNA based recombination in *Physcomitrella patens*). However, since the IMI gene normally differs from the AHAS gene at very few amino acids, a flanking sequence is not always necessary. The homologous recombination vector is introduced into a microorganism or plant cell (e.g., via polyethylene glycol mediated DNA), and cells in which the introduced IMI gene has homologously recombined with the endogenous AHAS gene are selected using art-known techniques.

In another embodiment, recombinant microorganisms can be produced that contain selected systems that allow for regulated expression of the introduced gene. For example, inclusion of an IMI gene on a vector placing it under control of the lac operon permits expression of the IMI gene only in the presence of IPTG. Such regulatory systems are well known in the art.

Whether present in an extra-chromosomal non-replicating vector or a vector that is integrated into a chromosome, the IMI polynucleotide preferably resides in a plant expression cassette. A plant expression cassette preferably contains regulatory sequences capable of driving gene expression in plant cells that are operably linked so that each sequence can fulfill its function, for example, termination of transcription by polyadenylation signals. Preferred polyadenylation signals are those originating from *Agrobacterium tumefaciens* t-DNA such as the gene 3 known as octopine synthase of the Ti-plasmid pTiACH5 (Gielen et al., 1984 EMBO J.

3:835) or functional equivalents thereof, but also all other terminators functionally active in plants are suitable. As plant gene expression is very often not limited on transcriptional levels, a plant expression cassette preferably contains other operably linked sequences like translational enhancers such as the overdrive-sequence containing the 5′-untranslated leader sequence from tobacco mosaic virus enhancing the polypeptide per RNA ratio (Gallie *et al.*, 1987 Nucl. Acids Research 15:8693-8711). Examples of plant expression vectors include those detailed in: Becker, D. *et al.*, 1992 New plant binary vectors with selectable markers located proximal to the left border, Plant Mol. Biol. 20:1195-1197; Bevan, M.W., 1984 Binary *Agrobacterium* vectors for plant transformation, Nucl. Acid. Res. 12:8711-8721; and Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Vol. 1, Engineering and Utilization, eds.: Kung and R. Wu, Academic Press, 1993, S. 15-38.

Plant gene expression should be operably linked to an appropriate promoter conferring gene expression in a timely, cell or tissue specific manner. Promoters useful in the expression cassettes of the invention include any promoter that is capable of initiating transcription in a plant cell. Such promoters include, but are not limited to those that can be obtained from plants, plant viruses and bacteria that contain genes that are expressed in plants, such as *Agrobacterium* and *Rhizobium*.

The promoter may be constitutive, inducible, developmental stage-preferred, cell type-preferred, tissue-preferred or organ-preferred. Constitutive promoters are active under most conditions. Examples of constitutive promoters include the CaMV 19S and 35 S promoters (Odell *et al.* 1985 Nature 313:810-812), the sX CaMV 35S promoter (Kay *et al.* 1987 Science 236:1299-1302) the Sep1 promoter, the rice actin promoter (McElroy *et al.* 1990 Plant Cell 2:163-171), the *Arabidopsis* actin promoter, the ubiquitan promoter (Christensen *et al.* 1989 Plant Molec Biol. 18:675-689); pEmu (Last *et al.* 1991 Theor Appl Genet. 81:581-588), the figwort mosaic virus 35S promoter, the Smas promoter (Velten *et al.* 1984 EMBO J. 3:2723-2730), the GRP1-8 promoter, the cinnamyl alcohol dehydrogenase promoter (U.S. Patent No. 5,683,439), promoters from the T-DNA of Agrobacterium, such as mannopine synthase, nopaline synthase, and octopine synthase, the small subunit of ribulose biphosphate carboxylase (ssuRUBISCO) promoter, and the like.

Inducible promoters are active under certain environmental conditions, such as the presence or absence of a nutrient or metabolite, heat or cold, light, pathogen attack, anaerobic conditions, and the like. For example, the hsp80 promoter from *Brassica* is induced by heat shock, the PPDK promoter is induced by light, the PR-1 promoter from tobacco, *Arabidopsis* and maize are inducible by infection with a pathogen, and the Adh1 promoter is induced by hypoxia and cold stress. Plant gene expression can also be facilitated via an inducible promoter (for review see Gatz, 1997 Annu. Rev. Plant Physiol. Plant Mol. Biol. 48:89-108). Chemically inducible promoters are especially suitable if gene expression is wanted to occur in a time specific manner. Examples of such promoters are a salicylic acid inducible promoter (PCT Application No. WO 95/19443), a tetracycline inducible promoter (Gatz et al., 1992 Plant J. 2:397-404) and an ethanol inducible promoter (PCT Application No. WO 93/21334).

Developmental stage-preferred promoters are preferentially expressed at certain stages of development. Tissue and organ preferred promoters include those that are preferentially expressed in certain tissues or organs, such as leaves, roots, seeds, or xylem. Examples of tissue preferred and organ preferred promoters include, but are not limited to fruit-preferred, ovule-preferred, male tissue-preferred, seed-preferred, integument-preferred, tuber-preferred, stalk-preferred, pericarp-preferred, and leaf-preferred, stigma-preferred, pollen-preferred, anther-preferred, a petal-preferred, sepal-preferred, pedicel-preferred, silique-preferred, stem-preferred, root-preferred promoters and the like. Seed preferred promoters are preferentially expressed during seed development and/or germination. For example, seed preferred promoters can be embryo-preferred, endosperm preferred and seed coat-preferred. See Thompson *et al.* 1989 BioEssays 10:108. Examples of seed preferred promoters include, but are not limited to cellulose synthase (celA), Cim1, gamma-zein, globulin-1, maize 19 kD zein (cZ19B1) and the like.

Other suitable tissue-preferred or organ-preferred promoters include the napin-gene promoter from rapeseed (U.S. Patent No. 5,608,152), the USP-promoter from *Vicia faba* (Baeumlein *et al.*, 1991 Mol Gen Genet. 225(3):459-67), the oleosin-promoter from *Arabidopsis* (PCT Application No. WO 98/45461), the phaseolin-promoter from *Phaseolus vulgaris* (U.S. Patent No. 5,504,200), the Bce4-promoter from Brassica (PCT Application No. WO 91/13980) or the legumin B4

promoter (LeB4; Baeumlein *et al.*, 1992 Plant Journal, 2(2):233-9) as well as promoters conferring seed specific expression in monocot plants like maize, barley, wheat, rye, rice, etc. Suitable promoters to note are the lpt2 or lpt1-gene promoter from barley (PCT Application No. WO 95/15389 and PCT Application No. WO 95/23230) or those described in PCT Application No. WO 99/16890 (promoters from the barley hordein-gene, rice glutelin gene, rice oryzin gene, rice prolamin gene, wheat gliadin gene, wheat glutelin gene, oat glutelin gene, Sorghum kasirin-gene and rye secalin gene).

Other promoters useful in the expression cassettes of the invention include, but are not limited to, the major chlorophyll a/b binding protein promoter, histone promoters, the Ap3 promoter, the -conglycin promoter, the napin promoter, the soy bean lectin promoter, the maize 15kD zein promoter, the 22kD zein promoter, the 27kD zein promoter, the g-zein promoter, the waxy, shrunken 1, shrunken 2 and bronze promoters, the Zm13 promoter (U.S. Patent No. 5,086,169), the maize polygalacturonase promoters (PG) (U.S. Patent Nos. 5,412,085 and 5,545,546) and the SGB6 promoter (U.S. Patent No. 5,470,359), as well as synthetic or other natural promoters.

Additional flexibility in controlling heterologous gene expression in plants may be obtained by using DNA binding domains and response elements from heterologous sources (i.e., DNA binding domains from non-plant sources). An example of such a heterologous DNA binding domain is the LexA DNA binding domain (Brent and Ptashne, Cell 43:729-736 (1985)).

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but they also apply to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell can be any prokaryotic or eukaryotic cell. For example, an IMI polynucleotide can be expressed in bacterial cells such as *C. glutamicum*, insect cells, fungal cells or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells), algae, ciliates, plant cells, fungi or other

microorganisms like *C. glutamicum*. Other suitable host cells are known to those skilled in the art.

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) an IMI polynucleotide. Accordingly, the invention further provides methods for producing IMI polypeptides using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding an IMI polypeptide has been introduced, or into which genome has been introduced a gene encoding a wild-type or IMI polypeptide) in a suitable medium until IMI polypeptide is produced. In another embodiment, the method further comprises isolating IMI polypeptides from the medium or the host cell. Another aspect of the invention pertains to isolated IMI polypeptides, and biologically active portions thereof. An "isolated" or "purified" polypeptide or biologically active portion thereof is free of some of the cellular material when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of IMI polypeptide in which the polypeptide is separated from some of the cellular components of the cells in which it is naturally or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of an IMI polypeptide having less than about 30% (by dry weight) of non-IMI material (also referred to herein as a "contaminating polypeptide"), more preferably less than about 20% of non-IMI material, still more preferably less than about 10% of non-IMI material, and most preferably less than about 5% non-IMI material.

When the IMI polypeptide, or biologically active portion thereof, is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the polypeptide preparation. The language "substantially free of chemical precursors or other chemicals" includes preparations of IMI polypeptide in which the polypeptide is separated from chemical precursors or other chemicals that are involved in the synthesis of the polypeptide. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of an IMI polypeptide

having less than about 30% (by dry weight) of chemical precursors or non-IMI chemicals, more preferably less than about 20% chemical precursors or non-IMI chemicals, still more preferably less than about 10% chemical precursors or non-IMI chemicals, and most preferably less than about 5% chemical precursors or non-IMI chemicals. In preferred embodiments, isolated polypeptides, or biologically active portions thereof, lack contaminating polypeptides from the same organism from which the IMI polypeptide is derived. Typically, such polypeptides are produced by recombinant expression of, for example, a *Triticum aestivum* IMI polypeptide in plants other than *Triticum aestivum* or microorganisms such as *C. glutamicum,* ciliates, algae or fungi.

The IMI polynucleotide and polypeptide sequences of the invention have a variety of uses. The nucleic acid and amino acid sequences of the present invention can be used to transform plants, thereby modulating the plant's resistance to imidazolinone herbicides. Accordingly, the invention provides a method of producing a transgenic plant having increased tolerance to an imidazolinone herbicide comprising, (a) transforming a plant cell with one or more expression vectors comprising one or more IMI nucleic acids, and (b) generating from the plant cell a transgenic plant with an increased resistance to an imidazolinone herbicide as compared to a wild type variety of the plant. In one embodiment, the multiple IMI nucleic acids are derived from different genomes. Also included in the present invention are methods of producing a transgenic plant having increased tolerance to an imidazolinone herbicide comprising, (a) transforming a plant cell with an expression vector comprising an IMI nucleic acid, wherein the nucleic acid is a non-Imi1 nucleic acid and (b) generating from the plant cell a transgenic plant with an increased resistance to an imidazolinone herbicide as compared to a wild type variety of the plant. .

The present invention includes methods of modifying a plant's tolerance to an imidazolinone herbicide comprising modifying the expression of one or more IMI nucleic acids. Preferably, the nucleic acids are located on or derived from different genomes. The plant's resistance to the imidazolinone herbicide can be increased or decreased as achieved by increasing or decreasing the expression of an IMI polynucleotide, respectively. Preferably, the plant's resistance to the imidazolinone herbicide is increased by increasing expression of an IMI

polynucleotide. Expression of an IMI polynucleotide can be modified by any method known to those of skill in the art. The methods of increasing expression of IMI polynucleotides can be used wherein the plant is either transgenic or not transgenic. In cases when the plant is transgenic, the plant can be transformed with a vector containing any of the above described IMI coding nucleic acids, or the plant can be transformed with a promoter that directs expression of endogenous IMI polynucleotides in the plant, for example. The invention provides that such a promoter can be tissue specific or developmentally regulated. Alternatively, non-transgenic plants can have endogenous IMI polynucleotide expression modified by inducing a native promoter. The expression of polynucleotides comprising SEQ ID NO:1 or SEQ ID NO:3 in target plants can be accomplished by, but is not limited to, one of the following examples: (a) constitutive promoter, (b) chemical-induced promoter, and (c) engineered promoter over-expression with for example zinc-finger derived transcription factors (Greisman and Pabo, 1997 Science 275:657).

In a preferred embodiment, transcription of the IMI polynucleotide is modulated using zinc-finger derived transcription factors (ZFPs) as described in Greisman and Pabo, 1997 Science 275:657 and manufactured by Sangamo Biosciences, Inc. These ZFPs comprise both a DNA recognition domain and a functional domain that causes activation or repression of a target nucleic acid such as an IMI nucleic acid. Therefore, activating and repressing ZFPs can be created that specifically recognize the IMI polynucleotide promoters described above and used to increase or decrease IMI polynucleotide expression in a plant, thereby modulating the herbicide resistance of the plant.

As described in more detail above, the plants produced by the methods of the present invention can be monocots or dicots. The plants can be selected from maize, wheat, rye, oat, triticale, rice, barley, soybean, peanut, cotton, rapeseed, canola, manihot, pepper, sunflower, tagetes, solanaceous plants, potato, tobacco, eggplant, tomato, Vicia species, pea, alfalfa, coffee, cacao, tea, Salix species, oil palm, coconut, perennial grass and forage crops, for example. Forage crops include, but are not limited to, Wheatgrass, Canarygrass, Bromegrass, Wildrye Grass, Bluegrass, Orchardgrass, Alfalfa, Salfoin, Birdsfoot Trefoil, Alsike Clover, Red Clover and Sweet Clover. In a preferred embodiment, the plant is a wheat plant. In each of the methods described above, the plant cell includes, but is not limited to, a

protoplast, gamete producing cell, and a cell that regenerates into a whole plant. As used herein, the term "transgenic" refers to any plant, plant cell, callus, plant tissue or plant part, that contains all or part of at least one recombinant polynucleotide. In many cases, all or part of the recombinant polynucleotide is stably integrated into a chromosome or stable extra-chromosomal element, so that it is passed on to successive generations.

As described above, the present invention teaches compositions and methods for increasing the imidazolinone resistance of a wheat plant or seed as compared to a wild-type variety of the plant or seed. In a preferred embodiment, the imidazolinone resistance of a wheat plant or seed is increased such that the plant or seed can withstand an imidazolinone herbicide application of preferably approximately 10-400 g ai ha^{-1}, more preferably 20-160 g ai ha^{-1}, and most preferably 40-80 g ai ha^{-1}. As used herein, to "withstand" an imidazolinone herbicide application means that the plant is either not killed or not injured by such application.

Additionally provided herein is a method of controlling weeds within the vicinity of a wheat plant, comprising applying an imidazolinone herbicide to the weeds and to the wheat plant, wherein the wheat plant has increased resistance to the imidazolinone herbicide as compared to a wild type variety of the wheat plant, and wherein the plant comprises one or more IMI nucleic acids. In one embodiment, the plant comprises multiple IMI nucleic acids located on or derived from different genomes. In another embodiment, the plant comprises a non-Imi1 nucleic acid. By providing for wheat plants having increased resistance to imidazolinone, a wide variety of formulations can be employed for protecting wheat plants from weeds, so as to enhance plant growth and reduce competition for nutrients. An imidazolinone herbicide can be used by itself for pre-emergence, post-emergence, pre-planting and at-planting control of weeds in areas surrounding the wheat plants described herein or an imidazolinone herbicide formulation can be used that contains other additives. The imidazolinone herbicide can also be used as a seed treatment. Additives found in an imidazolinone herbicide formulation include other herbicides, detergents, adjuvants, spreading agents, sticking agents, stabilizing agents, or the like. The imidazolinone herbicide formulation can be a wet or dry preparation and can include, but is not limited to, flowable powders, emulsifiable concentrates and liquid concentrates. The imidazolinone herbicide and herbicide formulations can be applied

in accordance with conventional methods, for example, by spraying, irrigation, dusting, or the like.

Throughout this application, various publications are referenced. The disclosures of all of these publications and those references cited within those publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains.

It should also be understood that the foregoing relates to preferred embodiments of the present invention and that numerous changes may be made therein without departing from the scope of the invention. The invention is further illustrated by the following examples, which are not to be construed in any way as imposing limitations upon the scope thereof. On the contrary, it is to be clearly understood that resort may be had to various other embodiments, modifications, and equivalents thereof, which, after reading the description herein, may suggest themselves to those skilled in the art without departing from the spirit of the present invention and/or the scope of the appended claims.

EXAMPLES

EXAMPLE 1

Mutagenesis and Selection of Resistant Wheat Lines

Approximately 40,000 seeds of *Triticum aestivum* L. cv CDC Teal (Hughes and Hucl, 1993 Can. J. Plant Sci. 73:193-197) were mutagenized using modified procedures described by Washington and Sears (1970). Seeds were pre-soaked in distilled water for four hours, followed by treatment with 0.3% EMS for six hours. Seeds were rinsed continually with tap water for seven hours and allowed to dry for approximately four hours before being planted in the field. The M_1 plants were selfed and the seed was harvested in bulk. Approximately 2 x 10^6 M_2 plants were grown in the field the following year and were sprayed at the two leaf stage with imazamox at a rate of 40 g ai ha^{-1} in a spray volume of 100 L ha^{-1}. Merge 0.05% (v/v) adjuvant was added to the spray solution. Six lines resistant to imazamox were selected and designated as lines 1A, 9A, 10A, 11A, 15A, and 16A. The M_3 and M_4 generations were grown in a walk-in growth chamber and plants resistant to

imazamox were selected using rates of 20 g ai ha^{-1}. Resistant plants were selected in the M_5 generation after application of 40 g ai ha^{-1} in the field. M_5 seed was homozygous for the trait, as progeny testing detected no segregation for resistance to imazamox.

EXAMPLE 2

Methods Used to Determine Inheritance and Allelism of IMI genes

To determine the genetic control of resistance to imazamox in the six wheat lines, reciprocal crosses between the six homozygous resistant M_6 lines and CDC Teal (susceptible to imazamox) were made. Randomly selected F_1 plants from each of the crosses were backcrossed to CDC Teal to form backcross (BC)F_1 populations. To investigate allelism, all possible inter-crosses between the six mutants and SWP965001 (Grandin/3*Fidel--FS-4) were made. SWP965001 is a spring wheat line that is homozygous for the FS-4 allele. Parental genotypes were grown in a walk-in growth chamber with a 16 hour photoperiod and a 24°C day and 16°C night temperature regime. Spikes that were ¾ emerged from the boot were emasculated and then pollinated 2–3 days after the emasculation date. Randomly selected F_2 plants from all segregating crosses were selfed to produce $F_{2:3}$ families. Parental, F_1, BCF_1, F_2 plants and $F_{2:3}$ families were tested for reaction to imazamox. All experiments were conducted in a walk-in growth chamber with a 16 hour photoperiod and a 23°C day and 16°C night temperature regime. A completely random design was used for all experiments. In experiments involving $F_{2:3}$ families, effort was taken to randomize both within and among families. The F_1 and F_2 populations were screened in the same experiment along with parental genotypes and CDC Teal as controls. Both the BCF_1 and $F_{2:3}$ populations were screened in two separate experiments along with appropriate parental genotypes as controls.

Herbicide treatments were applied to plants growing in 8 x 16 cell flats at the two leaf stage using a traveling cable sprayer calibrated to spray 100 L ha^{-1}. Imazamox was applied to plants at a rate of 20 g ai ha^{-1} using an 8001 EVS nozzle at a pressure of 275 kPa. Merge surfactant (0.05% v/v) was added to the herbicide solution prior to application. Fifteen days after herbicide application, plants were rated based on parental reactions and were considered as resistant, intermediate, or susceptible. Resistant plants were phenotypically unaffected following herbicide

treatment whereas intermediate plants were characterized by halted growth of the first two leaves, darkening (dark-green pigmentation) of the leaves, and the emergence of coleoptilar tillers. Susceptible plants were characterized by failure to develop new leaves, extensive leaf chlorosis, and eventually, plant death. For Mendelian analysis of the segregating populations, plants were scored into resistant and susceptible categories and tested for goodness of fit to various 1 gene, 2 gene and 3 gene models using chi-square analysis. For F_2 and BCF_1 plant data, intermediate reactions were included in the resistant reaction category. Yates correction for continuity was used to adjust the chi-square value when only a single degree of freedom was used in the chi-square analysis (Steele and Torrie 1980 Principles and procedures of statistics. McGraw-Hill, New York, New York. pp 633).

EXAMPLE 3

Results Regarding Inheritance of IMI genes

All resistant parents produced a similar phenotype when sprayed with 20 g ai ha^{-1} of imazamox. (Figure 1). Reciprocal crosses between the resistant lines and the susceptible parent (CDC Teal) resulted in F_1 plants that survived application of imazamox (Figure 1), indicating that resistance to imazamox is a nuclear and not a cytoplasmic trait. With the exception of cross 15A x Teal, the F_1 plants resulting from each of the resistant lines crossed with CDC Teal displayed an intermediate reaction (Figure 1). Since the F_1 plants were phenotypically intermediate between the two parents, it was concluded that resistance to imazamox in these lines was a partially dominant trait (Figure 1). Genetic analysis of resistance to imidazolinones and sulfonylureas in *Arabidopsis thaliana* (Haughn and Somerville, 1986 Mol. Gen. Genet. 204:430-434) *Zea mays* (Newhouse et al., 1991 Theor. Appl. Genet. 83:65-70), *Brassica napus* (Swanson et al., 1989 Theor. Appl. Gen. 78:525-530), and *Glycine max* (Sebastian et al., 1989 Crop Sci. 29:1403-1408) also indicated the presence of a single, partially dominant nuclear gene.

Fourteen F_1 plants resulting from the 15A x Teal cross were rated as resistant (Figure 1). Evaluation of F_2 populations from this cross indicated that two independently segregating loci were involved in conferring resistance in this genotype (Figure 2). Since the F_1 would carry two heterozygous resistant loci, one would expect that a resistant reaction would be observed. If each of these loci alone would

confer partial dominance, additively, two heterozygous loci would produce a resistant reaction. Swanson et al. (1989) combined two semi-dominant imidazolinone resistance alleles from *Brassica napus*, representing two unlinked genes, to produce a F_1 hybrid that was superior in imidazolinone resistance than either of the heterozygous lines alone. The authors concluded that resistance mechanisms are additive, and a higher level of resistance is observed in lines carrying more than one resistance allele.

An analysis of cytoplasmic inheritance was conducted in the F_2 generation by testing homogeneity of deviations from segregation ratios between the two reciprocal F_2 populations. Chi-square analysis revealed no significant deviations between reciprocal populations, confirming the absence of cytoplasmic inheritance (Figure 2). Since cytoplasmic inheritance was absent, data from the two reciprocal populations was combined and a total chi-square on pooled F_2 data was calculated (Figure 2).

With the exception of Teal x 15A, all F_2 populations resulting from resistant x susceptible crosses gave a good fit to a 3:1 resistant susceptible ratio indicating segregation of a single major gene for resistance to imazamox (Figure 2). When F_1 plants were crossed to the susceptible parent, resulting BCF_1 populations gave a good fit to a 1:1 resistant:susceptible ratio, confirming the single locus hypothesis (Figure 2). The F_2 population data from the cross 15A x Teal fit a 15:1 resistant:susceptible ratio (P=0.08), indicating segregation of two independent, complementary genes (Figure 2). The BCF_1 population gave good fit to a 3:1 resistant:susceptible ratio with a chi-square P value of 0.35, confirming the results of the F_2 (Figure 2).

Since it is speculated from F_2 data that resistance in lines 1A, 9A, 10A, 11A, and 16A are controlled by a single major gene, $F_{2:3}$ families should segregate and fit a 1:2:1 homozygous resistance: segregating: homozygous susceptible family ratio. Evaluation of $F_{2:3}$ families indicated that crosses Teal x 1A, Teal x 9A, Teal x 10A, Teal x 11A, and Teal x 16A all fit a 1:2:1 resistant: segregating: susceptible $F_{2:3}$ family ratio with chi-square P values of 0.64, 0.66, 0.52, 0.40, and 0.94, respectively (Figure 3). These results confirm the results of the F_2 and BCF_1 data that resistance in lines 1A, 10A, 9A, 11A, and resistance in 16A is controlled by a single major gene. This pattern of inheritance is consistent with other findings that have reported the

genetic control of resistance to AHAS inhibitor herbicides. To date, nearly all plant mutations conferring resistance to imidazolinones show that a single, partially dominant gene controls the resistance trait. In *Triticum aestivum*, *Zea mays*, *Glycine max*, *Arabidopsis thaliana*, and *Nicotiana tabacum*, resistance to AHAS inhibitors is inherited as a single partially dominant nuclear gene (Newhouse et al. 1991; Newhouse et al. 1992; Chaleff and Ray, 1984 Science 223:1148-1151; Sathasivan et al., 1991 Plant Physiol. 97:1044-1050). Plant resistance to AHAS inhibitor herbicides occurs mostly because of a single point mutation to the gene encoding the AHAS enzyme (Harms et al. 1992, Mol. Gen. Genet. 233:427-435; Winder and Spalding, 1988 Mol. Gen. Genet. 238:394-399).

The F_2 data resulting from the cross Teal x 15A provided a good fit to a 15:1 resistant:susceptible ratio, suggesting segregation of two, independently segregating loci (Figure 2). If this is the case, $F_{2:3}$ families should segregate and fit a 7:8:1 resistant:segregating:susceptible $F_{2:3}$ family ratio. $F_{2:3}$ families from the cross 15A x Teal did fit the expected 7:8:1 ratio (Figure 3), confirming the results of the F_2 and BCF_1 populations that resistance in 15A is conferred by two, independent loci. To the inventor's knowledge, this is the first reported instance were two independently segregating imidazolinone resistant alleles were identified in a single line following seed mutagenesis.

EXAMPLE 4

Results Regarding Allelism of IMI genes

To determine the allelic relationships of resistance genes, all possible intercrosses between resistant lines were evaluated. No susceptible plants were observed in the F_2 populations resulting from the inter-crosses between lines SWP965001, 1A, 9A, 10A, 15A, and 16A (Figure 4). Since these populations were not segregating, the resistance genes in these lines are either alleles at the FS-4 locus, or are very tightly linked. Since these populations were not segregating in the F_2 generation, $F_{2:3}$ families from these crosses were not evaluated.

All inter-crosses involving line 11A did segregate in the F_2 generation, indicating the presence of a unique resistance gene in 11A (Figure 4). If two independently segregating resistance genes are present as the result of crossing two lines, each carrying a single resistance gene, a 15:1 resistant:susceptible ratio would

be expected in the F_2 generation. In the F_2 generation, crosses SWP965001 x 11A, 1A x 11A, 10A x 11A, and 16A x 11A fit the expected 15:1 resistant:susceptible ratio suggesting independent segregation of two major resistance genes (Figure 4). $F_{2:3}$ family ratios from these crosses also gave a good fit to a 7:8:1 resistant:segregating:susceptible ratio, confirming the results obtained in the F_2 generation (Figure 5). Cross 11A x 9A did produce a segregating F_2 population, but the ratio did not fit a 15:1 segregation ratio due to an excess of susceptible segregants. Various other two gene hypotheses were tested, but all were found to be highly significant (Data not shown). Evaluation of $F_{2:3}$ families from this cross, however did give good fit to a 7:8:1 segregation ratio, indicating segregation of two independent genes (Figure 5). These results confirm that the resistance gene in 11A is unique from those in lines SWP96001, 1A, 9A, 10A, and 16A.

Cross 11A x 15A did produce a segregating F_2 population. Since 15A is carrying two resistance genes, one allelic to FS-4, a segregating F_2 population in cross 11A x 15A would indicate the presence of three segregating genes. Segregating generations resulting from cross 15A x 11A were tested for segregation of three independent loci. F_2 plants did fit the expected 63:1 resistant:susceptible ratio, indicating the segregation of three independent loci (Figure 4). These results suggest that the second mutation in 15A is not allelic to the resistance gene in 11A. $F_{2:3}$ families were not screened as over 330 plants within each family would have to be screened in order to ensure an adequate power of test (Hanson, 1959 Agron. J. 51:711-716).

Three independent resistance loci have been identified, each with an allele conferring resistance to imazamox. Recommended rules for gene locus and allele symbolization have been published (McIntosh et al., 1998 Catalogue of Gene Symbols. Volume 5, Proceedings of the 9th International Wheat Genetics Symposium. Saskatoon, Saskatchewan). Non-allelic gene loci of an enzyme that catalyze the same reaction should be given the same symbol, corresponding to the trivial name of the enzyme. The trivial name for AHAS is ALS. Absent data to assign the loci to specific chromosomes and genomes, they should be designated in sequential series. The designation of the phenotype observed when changes occur in the gene resulting in a new allele should reflect that phenotype. Thus, it is proposed that the FS-4 imidazolinone resistance allele be designated as *Imi1* and the locus it is

at designated as *Als1*. *Imi* stands for imidazolinone resistance. This designation indicates that the gene is a dominant trait and it is the first allele identified. Segregating F_2 and $F_{2:3}$ population data suggests that 15A and 11A carry two new independent resistance alleles at different loci (Figures 2 and 3). The designations for these alleles are *Imi2* for the 11A mutation at the *Als2* locus and *Imi3* for the second 15A mutation at the *Als3* locus.

Identified herein are three independently segregating alleles conferring resistance to imazamox, namely *Imi1*(1A, 9A, 10A, 15A and 16A), *Imi2* (11A), and *Imi3* (15A). It is proposed that each of the three identified alleles are associated with a different structural gene coding for herbicide-insensitive forms of AHAS. Since wheat is a hexaploid, multiple AHAS loci would be expected. Other polyploid species have been found to have more than one copy of AHAS. In *Nicotiana tabacum*, an allotetraploid, two AHAS genes have been identified and characterized (Mazur et al. 1987). Chaleff and Ray (1984) identified two independently segregating sulfonylurea resistance alleles in *Nicotiana tabacum*, each coding for an altered form of AHAS. *Zea mays* possesses two constitutively expressed identical AHAS genes (Fang et al., 1992 Plant Mol. Biol. 18:1185-1187). In allotetraploid *Brassica napus* and *Gossypium hirsutum*, an AHAS multi-gene family consisting of five and six members, respectively, is present (Rutledge et al., 1991 Mol Gen. Genet. 229:31-40; Grula et al., 1995 Plant Mol. Biol. 28:837-846). Higher levels of resistance to herbicides have been observed in polyploid species when multiple resistance alleles are present. Swanson et al. (1989 Theor. Appl. Gen. 78:525-530) combined two unique imidazolinone resistance alleles from two homozygous *Brassica napus* lines resulting in progeny with a higher level of resistance than either homozygous line alone. Creason and Chaleff (1988 Theor. Appl. Genet. 76:177-182) identified *Nicotiana tabacum* plants homozygous for two mutations that conferred resistance to sulfonylureas. Plants homozygous for both mutations were five-fold more resistant to foliar applications of chlorsulfuron than were plants homozygous for each single mutation. The present invention proposes producing increased levels of resistance to an imidazolinone herbicide in wheat by combining any two or all three resistance alleles.

EXAMPLE 5

Tolerance to Imidazolinone Herbicides in Teal11A, Teal15A and Teal11A/15A Hybrid

The increased tolerance exhibited by Teal 11A and Teal 15A to 20 grams per hectare of imazamox has been exemplified in previous examples by the ability to distinguish tolerant from susceptible parental and segregant plants in inheritance studies. Teal 11A has been shown to confer similar levels of tolerance to imidazolinone herbicides to that conferred by the FS4 mutation in Fidel in various greenhouse and field comparisons. The similarity in tolerance is also reflected in comparing the *in vitro* activity of AHAS extracted from tolerant plants. This is possible because the tolerance in Teal 11A, Teal 15A, and FS4 is due to a mutation in the AHAS enzyme rendering it resistant to inhibition by imidazolinone herbicides. Figure 6 indicates that the activity of AHAS enzyme extracted from Teal 11A and BW755, a line containing FS4, changes similarly as the rate of imazamox increases, and both have a higher percentage of active (resistant) enzyme at the highest concentration of imazamox than does the wild type check, Teal.

The presence of two IMI nucleic acids in Teal 15A provides increased tolerance to imidazolinone herbicides compared to a line such as BW755 carrying only one IMI nucleic acid. This increased tolerance is reflected both in less injury at higher herbicide rates, but in having more uninhibited AHAS enzyme activity. Figure 7 illustrates that a 10x rate of imazamox (200 g/ha), all treated one gene plants were injured, while no two gene plants were injured. At all concentrations of imazamox in an *in vitro* assay of AHAS activity (Figure 6), but particularly at the highest concentrations, Teal15A had a higher percentage of active (resistant) enzyme than did either of the single gene lines, Teal11A and BW755.

Combining three non-allelic genes each conferring tolerance to imidazolinone herbicides results in greater tolerance than with only two non-allelic genes (Figure 7). At a 30X rate, or 600 g/ha of imazamox, over half of plants sustained no injury in a still-segregating selfed population of Teal15A crossed with Teal11A, while all plants of the homozygous population of Teal15A sustained injury.

CLAIMS

We Claim:

1. A wheat plant comprising multiple IMI nucleic acids, wherein the nucleic acids are from different genomes and wherein the wheat plant has increased resistance to an imidazolinone herbicide as compared to a wild-type variety of the plant.

2. The wheat plant of claim 1, wherein the multiple IMI nucleic acids are selected from the group consisting of an Imi1 nucleic acid, an Imi2 nucleic acid and an Imi3 nucleic acid.

3. The wheat plant of claim 1, wherein the multiple IMI nucleic acids encode proteins comprising a mutation in a conserved amino acid sequence selected from the group consisting of a Domain A, a Domain B, a Domain C, a Domain D and a Domain E.

4. The wheat plant of claim 3, wherein the conserved amino acid sequence is a Domain E.

5. The wheat plant of claim 4, wherein the mutation results in a serine to asparagine substitution in the IMI protein as compared to a wild-type AHAS protein.

6. The wheat plant of claim 1, wherein the multiple nucleic acids are selected from the group consisting of:

   a) a polynucleotide comprising SEQ ID NO:1;
   b) a polynucleotide comprising SEQ ID NO:3;
   c) a polynucleotide encoding a polypeptide comprising SEQ ID NO:2;
   d) a polynucleotide encoding a polypeptide comprising SEQ ID NO:4;
   e) a polynucleotide comprising at least 60 consecutive nucleotides of any of a) through d); and

f) a polynucleotide complementary to the polynucleotide of any of a) through e).

7. The wheat plant of claim 1, wherein one of the IMI nucleic acids comprises a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2.

8. The wheat plant of claim 1, wherein one of the IMI nucleic acids comprises a polynucleotide sequence encoding the polypeptide of SEQ ID NO:4.

9. The wheat plant of claim 1, comprising two IMI nucleic acids.

10. The wheat plant of claim 9, comprising an Imi1 nucleic acid and an Imi3 nucleic acid.

11. The wheat plant of claim 9, comprising an Imi1 nucleic acid and an Imi2 nucleic acid.

12. The wheat plant of claim 10 or claim 11, wherein the Imi1 nucleic acid comprises a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2.

13. The wheat plant of claim 1, comprising three IMI nucleic acids.

14. The wheat plant of any one of claims 1-5, wherein the plant is transgenic.

15. The wheat plant of any one of claims 1-5, wherein the plant is not transgenic.

16. The wheat plant of claim 15, wherein the plant has an ATCC Patent Deposit Designation Number PTA-3955, PTA-3954 or PTA-3953; or is a recombinant or genetically engineered derivative of the plant with ATCC Patent Deposit Designation Number PTA-3955, PTA-3954 or PTA-3953; or of any progeny of the plant with ATCC Patent Deposit Designation Number PTA-3955, PTA-3954 or PTA-3953; or is a plant that is a progeny of any of these plants.

17. The wheat plant of claim 15, wherein the plant has an ATCC Patent Deposit Designation Number PTA-3955, PTA-3954 or PTA-3953, or is a progeny of the plant with ATCC Patent Deposit Designation Number PTA-3955, PTA-3954 or PTA-3953.

18. The wheat plant of claim 15, wherein the plant has the herbicide resistance characteristics of the plant with ATCC Patent Deposit Designation Number PTA-3955, PTA-3954 or PTA-3953.

19. The wheat plant of claim 15, wherein the wheat plant has an ATCC Patent Deposit Designation Number PTA-3955, PTA-3954 or PTA-3953.

20. The wheat plant of claim 1, wherein the imidazolinone herbicide is selected from the group consisting of 2-(4-isopropyl-4-methyl-5-oxo-2-imidiazolin-2-yl)-nicotinic acid, 2-(4-isopropyl)-4-methyl-5-oxo-2-imidazolin-2-yl)-3-quinolinecarboxylic acid, 5-ethyl-2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-nicotinic acid, 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-5-(methoxymethyl)-nicotinic acid, 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-5-methylnicotinic acid, and a mixture of methyl 6-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-m-toluate and methyl 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-p-toluate.

21. The wheat plant of claim 1, wherein the imidazolinone herbicide is 5-ethyl-2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-nicotinic acid.

22. The wheat plant of claim 1, wherein the imidazolinone herbicide is 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-5-(methoxymethyl)-nicotinic acid.

23. A plant part of the wheat plant of claim 1.

24. A plant cell of the wheat plant of claim 1.

25. A seed produced by the wheat plant of claim 1.

26. The seed of claim 25, wherein the seed is true breeding for an increased resistance to an imidazolinone herbicide as compared to a wild type variety of the wheat plant seed.

27. A wheat plant comprising an IMI nucleic acid, wherein the nucleic acid is a non-Imi1 nucleic acid and wherein the wheat plant has increased resistance to an imidazolinone herbicide as compared to a wild-type variety of the plant.

28. The wheat plant of claim 27, wherein the IMI nucleic acid is an Imi2 nucleic acid.

29. The wheat plant of claim 27, wherein the IMI nucleic acid comprises a polynucleotide sequence encoding a polypeptide of SEQ ID NO:4.

30. The wheat plant of claim 27, wherein the IMI nucleic acid comprises a polynucleotide sequence of SEQ ID NO:3.

31. The wheat plant of claim 27, wherein the imidazolinone herbicide is selected from the group consisting of 2-(4-isopropyl-4-methyl-5-oxo-2-imidiazolin-2-yl)-nicotinic acid, 2-(4-isopropyl)-4-methyl-5-oxo-2-imidazolin-2-yl)-3-quinolinecarboxylic acid, 5-ethyl-2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-nicotinic acid, 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-5-(methoxymethyl)-nicotinic acid, 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-5-methylnicotinic acid, and a mixture of methyl 6-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-m-toluate and methyl 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-p-toluate.

32. The wheat plant of claim 27, wherein the imidazolinone herbicide is 5-ethyl-2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-nicotinic acid.

33. The wheat plant of claim 27, wherein the imidazolinone herbicide is 2-(4-isopropyl-4-methyl-5-oxo-2-imidazolin-2-yl)-5-(methoxymethyl)-nicotinic acid.

34. A plant part of the wheat plant of claim 27.

35. A plant cell of the wheat plant of claim 27.

36. A seed produced by the wheat plant of claim 27.

37. The seed of claim 36, wherein the seed is true breeding for an increased resistance to an imidazolinone herbicide as compared to a wild type variety of the wheat plant seed.

38. The wheat plant of any one of claims 27-30, wherein the plant is transgenic.

39. The wheat plant of any one of claims 27-30, wherein the plant is not transgenic.

40. The wheat plant of claim 39, wherein the plant has an ATCC Patent Deposit Designation Number PTA-3953; or is a recombinant or genetically engineered derivative of the plant with ATCC Patent Deposit Designation Number PTA-3953; or of any progeny of the plant with ATCC Patent Deposit Designation Number PTA-3953; or is a plant that is a progeny of any of these plants.

41. The wheat plant of claim 39, wherein the plant has an ATCC Patent Deposit Designation Number PTA-3953 or is a progeny of the plant with ATCC Patent Deposit Designation Number PTA-3953.

42. The wheat plant of claim 39, wherein the plant has the herbicide resistance characteristics of the plant with ATCC Patent Deposit Designation Number PTA-3953.

43. The wheat plant of claim 39, wherein the wheat plant has an ATCC Patent Deposit Designation Number PTA-3953.

44. An isolated IMI nucleic acid, wherein the nucleic acid comprises a polynucleotide selected from the group consisting of:

a) a polynucleotide of SEQ ID NO:1;

b) a polynucleotide of SEQ ID NO:3;

c) a polynucleotide encoding a polypeptide of SEQ ID NO:2;

d) a polynucleotide encoding a polypeptide of SEQ ID NO:4;

e) a polynucleotide comprising at least 60 consecutive nucleotides of any of a) through d); and

f) a polynucleotide complementary to the polynucleotide of any of a) through e).

45. The isolated IMI nucleic acid of claim 44, wherein the nucleic acid comprises a polynucleotide of SEQ ID NO:1.

46. The isolated IMI nucleic acid of claim 44, wherein the nucleic acid comprises a polynucleotide of SEQ ID NO:3.

47. The isolated IMI nucleic acid of claim 44, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide of SEQ ID NO:2.

48. The isolated IMI nucleic acid of claim 44, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide of SEQ ID NO:4.

49. A method of controlling weeds within the vicinity of a wheat plant, comprising applying an imidazolinone herbicide to the weeds and to the wheat plant, wherein the wheat plant has increased resistance to the imidazolinone herbicide as compared to a wild type variety of the wheat plant, wherein the plant comprises multiple IMI nucleic acids, and wherein the nucleic acids are from different genomes.

50. The method of claim 49, wherein the multiple IMI nucleic acids are selected from the group consisting of an Imi1 nucleic acid, an Imi2 nucleic acid and an Imi3 nucleic acid.

51. The method of Claim 50, wherein the plant comprises an Imi1 nucleic acid and an Imi2 nucleic acid.

52. The method of Claim 50, wherein the plant comprises an Imi1 nucleic acid and an Imi3 nucleic acid.

53. The method of claim 49, wherein the multiple nucleic acids are selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO:1;

b) a polynucleotide comprising SEQ ID NO:3;

c) a polynucleotide encoding a polypeptide comprising SEQ ID NO:2;

d) a polynucleotide encoding a polypeptide comprising SEQ ID NO:4;

e) a polynucleotide comprising at least 60 consecutive nucleotides of any of a) through d); and

f) a polynucleotide complementary to the polynucleotide of any of a) through e).

54. A method of controlling weeds within the vicinity of a wheat plant, comprising applying an imidazolinone herbicide to the weeds and to the wheat plant, wherein the wheat plant has increased resistance to the imidazolinone herbicide as compared to a wild type variety of the wheat plant, and wherein the plant comprises an IMI nucleic acid that is a non-Imi1 nucleic acid.

55. The method of claim 54, wherein the IMI nucleic acid is selected from the group consisting of an Imi2 nucleic acid and an Imi3 nucleic acid.

56. The method of claim 54, wherein the IMI nucleic acid is selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO:3;

b) a polynucleotide encoding a polypeptide comprising SEQ ID NO:4;

c) a polynucleotide comprising at least 60 consecutive nucleotides of any of a) through b); and

d) a polynucleotide complementary to the polynucleotide of any of a) through c).

57. A method of modifying a plant's tolerance to an imidazolinone herbicide comprising modifying the expression of multiple IMI nucleic acids, wherein the nucleic acids are from different genomes.

58. The method of claim 57, wherein the multiple IMI nucleic acids are selected from the group consisting of an Imi1 nucleic acid, an Imi2 nucleic acid and an Imi3 nucleic acid.

59. The method of Claim 58, wherein the plant comprises an Imi1 nucleic acid and an Imi2 nucleic acid.

60. The method of Claim 58, wherein the plant comprises an Imi1 nucleic acid and an Imi3 nucleic acid.

61. The method of claim 57, wherein the multiple nucleic acids are selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO:1;

b) a polynucleotide comprising SEQ ID NO:3;

c) a polynucleotide encoding a polypeptide comprising SEQ ID NO:2;

d) a polynucleotide encoding a polypeptide comprising SEQ ID NO:4;

e) a polynucleotide comprising at least 60 consecutive nucleotides of any of a) through d); and

f) a polynucleotide complementary to the polynucleotide of any of a) through e).

62. A method of modifying a plant's tolerance to an imidazolinone herbicide comprising modifying the expression of an IMI nucleic acid, wherein the nucleic acid is a non-Imi1 nucleic acid.

63. The method of claim 62, wherein the IMI nucleic acid is selected from the group consisting of an Imi2 nucleic acid and an Imi3 nucleic acid.

64. The method of claim 62, wherein the IMI nucleic acid is selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO:3;

b) a polynucleotide encoding a polypeptide comprising SEQ ID NO:4;

c) a polynucleotide comprising at least 60 consecutive nucleotides of any of a) through b); and

d) a polynucleotide complementary to the polynucleotide of any of a) through c).

65. A method of producing a transgenic plant having increased resistance to an imidazolinone herbicide comprising,

a) transforming a plant cell with one or more expression vectors comprising multiple IMI nucleic acids, wherein the nucleic acids are derived from different genomes; and

b) generating from the plant cell a transgenic plant with an increased resistance to an imidazolinone herbicide as compared to a wild type variety of the plant.

66. The method of claim 65, wherein the multiple IMI nucleic acids are selected from the group consisting of an Imi1 nucleic acid, an Imi2 nucleic acid and an Imi3 nucleic acid.

67. The method of Claim 66, wherein the plant comprises an Imi1 nucleic acid and an Imi2 nucleic acid.

68. The method of Claim 66, wherein the plant comprises an Imi1 nucleic acid and an Imi3 nucleic acid.

69. The method of claim 65, wherein the multiple nucleic acids are selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO:1;

b) a polynucleotide comprising SEQ ID NO:3;

c) a polynucleotide encoding a polypeptide comprising SEQ ID NO:2;

d) a polynucleotide encoding a polypeptide comprising SEQ ID NO:4;

e) a polynucleotide comprising at least 60 consecutive nucleotides of any of a) through d); and

f) a polynucleotide complementary to the polynucleotide of any of a) through e).

70. A method of producing a transgenic plant having increased resistance to an imidazolinone herbicide comprising,

a) transforming a plant cell with an expression vector comprising an IMI nucleic acid, wherein the nucleic acid is a non-Imi1 nucleic acid; and

b) generating from the plant cell a transgenic plant with an increased resistance to an imidazolinone herbicide as compared to a wild type variety of the plant.

71. The method of claim 70, wherein the IMI nucleic acid is selected from the group consisting of an Imi2 nucleic acid and an Imi3 nucleic acid.

72. The method of claim 70, wherein the IMI nucleic acid is selected from the group consisting of:

a) a polynucleotide comprising SEQ ID NO:3;

b) a polynucleotide encoding a polypeptide comprising SEQ ID NO:4;

c) a polynucleotide comprising at least 60 consecutive nucleotides of any of a) through b); and

d) a polynucleotide complementary to the polynucleotide of any of a) through c).

Figure 1

Cross		Resistant (R)	Intermediate (I)	Susceptible (S)	Total F_1 screened		
					R	I	S
Teal		0	0	334			
1A		59	0	0			
	Teal x 1A	0	5	0	0	10	0
	1A x Teal	0	5	0			
9A		66	0	0			
	Teal x 9A	0	5	0	0	10	0
	9A x Teal	0	5	0			
10A		66	0	0			
	Teal x 10A	0	5	0	0	11	0
	10A x Teal	0	6	0			
11A		53	0	0			
	Teal x 11A	0	7	0	0	15	0
	11A x Teal	0	8	0			
15A		48	0	0			
	Teal x 15A	7	0	0	14	0	0
	15A x Teal	7	0	0			
16A		66	0	0			
	Teal x 16A	0	7	0	0	14	0
	16A x Teal	0	7	0			

Figure 2

	F$_2$ Generation					BCF$_1$ Generation[c]			
Cross	Resistant (R)	Susceptible (S)	Ratio tested (R:S)	P Value[a]	Homogeneity of reciprocal crosses[b]	R	S	Ratio tested (R:S)[d]	P Value[e]
Teal x 1A	635	201	3:1	0.55	0.49	33	27	1:1	0.52
Teal x 9A	492	141	3:1	0.12	0.66	42	34	1:1	0.39
Teal x 10A	701	231	3:1	0.91	0.87	27	20	1:1	0.38
Teal x 11A	505	189	3:1	0.19	0.32	34	30	1:1	0.71
Teal x 15A	893	74	15:1	0.08	0.39	45	20	3:1	0.35
Teal x 16A	587	168	3:1	0.09	0.33	26	23	1:1	0.78

Figure 3

Cross	Resistant (R)	Segregating (Seg)	Susceptible (S)	Ratio Tested (R:Seg:S)[a]	P Value[b]	Total # Families
Teal x 1A	12	28	10	1:2:1	0.64	50
Teal x 9A	15	23	12	1:2:1	0.66	50
Teal x 10A	9	27	14	1:2:1	0.52	50
Teal x 11A	14	26	8	1:2:1	0.40	48
Teal x 15A	36	55	9	7:8:1	0.21	100
Teal x 16A	12	25	11	1:2:1	0.94	48

Figure 4

Cross	Resistant (R)	Susceptible (S)	Ratio tested (R:S)	P Value[a]
1A x 9A	506	0	-	-
1A x 10A	567	0	-	-
1A x 15A	501	0	-	-
1A x 16A	814	0	-	-
1A x SWP965001	309	0	-	-
9A x 10A	603	0	-	-
9A x 15A	424	0	-	-
9A x 16A	336	0	-	-
9A x SWP965001	407	0	-	-
10A x 15A	547	0	-	-
10A x 16A	409	0	-	-
10A x SWP965001	686	0	-	-
15A x 16A	298	0	-	-
15A x SWP965001	410	0	-	-
16A x SWP965001	509	0	-	-
SWP965001 x 11A	688	47	15:1	0.93
1A x 11A	735	56	15:1	0.37
11A x 9A	446	42	15:1	0.00
11A x 10A	557	46	15:1	0.19
11A x 15A	600	14	63:1	0.20
11A x 16A	390	34	15:1	0.16

Figure 5

Cross	Resistant (R)	Segregating (Seg)	Susceptible (S)	Ratio Tested (R:Seg:S)[a]	P Value[b]	Total # Families
SWP965001 x 11A	32	42	7	7:8:1	0.57	81
1A x 11A	33	58	9	7:8:1	0.07	100
11A x 9A	36	49	5	7:8:1	0.77	90
11A x 10A	34	59	7	7:8:1	0.14	100
11A x 16A	45	47	7	7:8:1	0.86	99

Figure 6

Imazamox Conc. (uM)	% Uninhibited AHAS Activity			
	Teal	11A	BW755	15A
0.8	71.1	84.8	90.0	92.6
1.6	55.2	71.2	76.9	88.4
3.1	41.8	59.7	65.7	84.3
6.3	30.9	50.3	56.6	80.3
12.5	22.3	42.9	49.6	76.3
25.0	16.2	37.6	44.5	72.5
50.0	12.6	34.3	41.4	68.8
100.0	11.4	33.1	40.4	65.2

Figure 7

	10X		30X	
Genotype	%Injured	%No Injury	%Injured	%No Injury
BW755 (one gene)	100	0	100	0
Teal 15A (two gene)	0	100	100	0
15A/11A (three gene)	33	67	48	52

Figure 8

```
                                       10                                                    20
---------------------------------------+-----------------------------------------------------+-------------------------   SEQ ID NO:
 M   A   T   T   A   A   A   A   A   A   A   L   S   A   A   A   T   A   K   T   G   R   K   N   H   Q   R     14
ATG GCT ACG ACC GCC GCG GCC GCG GCC GCC GCC CTG TCC GCC GCC GCG ACG GCC AAG ACC GGC CGT AAG AAC CAC CAG CGA    13

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     16
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    15

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     18
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    17

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     20
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    19

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     22
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    21

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     24
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    23

         30                                      40                                              50
---------+---------------------------------------+-----------------------------------------------+-------------
 H   H   V   L   P   A   R   G   R   V   G   A   A   A   V   R   C   S   A   V   S   P   V   T   P   P   S     14
CAC CAC GTC CTT CCC GCT CGA GGC CGG GTG GGG GCG GCG GCG GTC AGG TGC TCG GCG GTG TCC CCG GTC ACC CCG CCG TCC    13

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     16
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    15

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     18
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    17

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     20
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    19

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     22
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    21

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X     24
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---    23
```

Figure 8 Continued

```
                          60                                                  70                                                  80
--------------------------+---------------------------------------------------+---------------------------------------------------+-----
 P   A   P   P   A   T   P   L   R   P   W   G   P   A   E   P   R   K   G   A   D   I   L   V   E   A   L
CCG GCG CCG CCG GCC ACG CCG CTC CGG CCG TGG GGG CCG GCC GAG CCC CGC AAG GGC GCG GAC ATC CTC GTG GAG GCG CTG

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                                      90                                                 100
--------------------------------------+--------------------------------------------------+-----------------------------
 E   R   C   G   V   S   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
GAG CGG TGC GGC GTC AGC GAC GTG TTC GCC TAC CCG GGC GGC GCG TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCC CCG

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCC GGC GGC GCC TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCC

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCC GGC GGC GCC TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCC

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCC GGC GGC GCC TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCC

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCC GGC GGC GCC TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCC

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCC GGC GGC GCC TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCC
```

Figure 8 Continued

```
     110                                                120                                                130
-----+--------------------------------------------------+--------------------------------------------------+---------------------
 V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGC GAG GCG TTC GCG GCG TCC GGG TAC GCG CGC GCG TCC GGC CGC GTC

 V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGC TAC GCC CGC GCG TCC GGC CGC GTC

 V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGC TAC GCC CGC GCG TCC GGC CGC GTC

 V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGC TAC GCC CGC GCG TCC GGC CGC GTC

 V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGC TAC GCC CGC GCG TCC GGC CGC GTC

 V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGC TAC GCC CGC GCG TCC GGC CGC GTC

                 140                                                150                                                160
-----------------+--------------------------------------------------+--------------------------------------------------+---------
 G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   V   P
GGG GTC TGC GTC GCC ACC TCC GGC CCC GGG GCA ACC AAC CTC GTG TCC GCG CTC GCC GAC GCG CTG CTC GAC TCC GTC CCG

 G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCC CTC CTC GAC TCC ATC CCC

 G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCC CTC CTC GAC TCC ATC CCC

 G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCC CTC CTC GAC TCC ATC CCC

 G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCC CTC CTC GAC TCC ATC CCC

 G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCC CTC CTC GAC TCC ATC CCC
```

Figure 8 Continued

```
                              170                                         180
------------------------------+-------------------------------------------+------------------------------------
 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACC GAC GCC TTC CAG GAG ACG CCC ATA GTC GAG GTC ACC

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAC GCG TTC CAG GAG ACG CCC ATA GTG GAG GTC ACG

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAC GCG TTC CAG GAG ACG CCC ATA GTG GAG GTC ACG

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAC GCG TTC CAG GAG ACG CCC ATA GTG GAG GTC ACG

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAC GCG TTC CAG GAG ACG CCC ATA GTG GAG GTC ACG

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAC GCG TTC CAG GAG ACG CCC ATA GTG GAG GTC ACG

190                                     200                                     210
-+--------------------------------------+---------------------------------------+-------------------------
 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAT TAC CTT GTC CTT GAT GTG GAG GAC ATC CCC CGC GTC ATA CAG GAA GCC TTC TTC CTC GCG

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTT GCA

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTT GCA

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTT GCA

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTT GCA

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTT GCA
```

Figure 8 Continued

```
          220                                 230                                 240
----------+-----------------------------------+-----------------------------------+-----------
 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCG GGC CGT CCT GGC CCG GTG CTG GTC GAC ATC CCC AAG GAC ATC CAG CAG CAG ATG GCC GTG CCG GTC TGG GAC ACC

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTA GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCC GTC TGG GAC ACT

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTA GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCC GTC TGG GAC ACT

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTA GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCC GTC TGG GAC ACT

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTA GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCC GTC TGG GAC ACT

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTA GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCC GTC TGG GAC ACT

                      250                                 260                                 270
----------------------+-----------------------------------+-----------------------------------+-
 S   M   N   L   P   G   Y   I   A   R   L   P   K   P   P   A   T   E   L   L   E   Q   V   L   R   L   V
TCG ATG AAT CTA CCA GGG TAC ATC GCA CGC CTG CCC AAG CCA CCC GCG ACA GAA TTG CTT GAG CAG GTC TTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCA ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCA ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCA ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCA ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCA ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT
```

Figure 8 Continued

```
                                    280                                                     290
-------------------------------------+--------------------------------------------------------+----------------------------
 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   S   A   S   G   D   E   L   R   W   F   V   E   L
GGC GAG TCA CGG CGC CCG ATT CTC TAT GTC GGT GGT GGC TGC TCT GCA TCT GGT GAC GAA TTG CGC TGG TTT GTT GAG CTG

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCG TCT GGC GAG GAG TTG CGC CGC TTT GTT GAG CTT

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCG TCT GGC GAG GAG TTG CGC CGC TTT GTT GAG CTT

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCG TCT GGC GAG GAG TTG CGC CGC TTT GTT GAG CTT

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCG TCT GGC GAG GAG TTG CGC CGC TTT GTT GAG CTT

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCG TCT GGC GAG GAG TTG CGC CGC TTT GTT GAG CTT

        300                                       310                                                     320
---------+-----------------------------------------+--------------------------------------------------------+----------------
 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGT ATC CCA GTT ACA ACC ACT CTG ATG GGC CTC GGC AAT TTC CCC AGT GAC GAC CCG TTG TCC CTG CGC ATG CTT GGG

 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTG ATG GGC CTT GGC AAC TTC CC- AGC GAC GAC CCA CTG TCT CTG CGC ATG CTT GGG

 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTG ATG GGC CTT GGC AAC TTC CCC AGC GAC GAC CCA CTG TCT CTG CGC ATG CTT GGG

 T   G   I   T   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT NCA GTT ACA ACT ACT CTG ATG GGC CTT GGC AAC TTC CCC AGC GAC GAC CCA CTG TCT CTG CGC ATG CTT GGG

 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTG ATG GGC CTT GGC AAC TTC CCC AGC GAC GAC CCA CTG TCT CTG CGC ATG CTT GGG

 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTG ATG GGC CTT GGC AAC TTC CCC AGC GAC GAC CCA CTG TCT CTG CGC ATG CTT GGG
```

Figure 8 Continued

```
                              330                                               340                                               350
----------------------------+-------------------------------------------------+-------------------------------------------------+-----
 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   G   V   R   F   D   D   R
ATG CAT GGC ACG GTG TAC GCA AAT TAT GCC GTG GAT AAG GCT GAC CTG TTG CTT GCG TTT GGT GTG CGG TTT GAT GAT CGT

 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   G   V   R   F   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG CTC GCA TTT GGT GTG CGG TTT GAT GAT CGT

 M   N   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   G   V   R   F   D   D   R
ATG NNT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG CTC GCA TTT GGT GTG CGG TTT GAT GAT CGT

 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   D   V   R   F   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG CTC GCA TTT GNT GTG CGG TTT GAT GAT CGT

 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   I   T   F   G   V   R   F   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG NTC NCA TTT GGT GTG CGG TTT GAT GAT CGT

 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   I   A   F   G   V   R   F   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG NTC GCA TTT GGT GTG CGG TTT GAT GAT CGT


                                        360                                               370
----------------------------------------+-------------------------------------------------+------------------------------------
 V   T   G   K   I   E   A   F   A   S   R   A   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACA GGG AAA ATT GAG GCT TTT GCA AGC AGG GCC AAG ATT GTG CAC ATT GAC ATT GAT CCA GCA GAG ATT GGA AAG AAC

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACT GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GTG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACT GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GTG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACT GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GTG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAN AAC

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACT GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GTG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   E   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACT GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GNG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC
```

Figure 8 Continued

```
     380                                               390                                               400
-----+-------------------------------------------------+-------------------------------------------------+-------------------
 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   A   L   L   Q   Q   S   T   T
AAG CAA CCA CAT GTG TCA ATT TGC GCA GAT GTT AAG CTT GCT TTA CAG GGC TTG AAT GCT CTG CTA CAA CAG AGC ACA ACA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   D   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAG CTT GCT TTA CAG GGG TTG AAT GAT CTA TTA AAT GGG AGC AAA GCA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   D   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAG CTT GCT TTA CAG GGG TTG AAT GAT CTA TTA AAT GGG AGC AAA GCA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   D   L   L   N   G   S   K   A
AAN CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAG CTT GCT TTA CAG GGG TTG AAT GAT CTA TTA AAT GGG AGC AAA GCA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   D   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCC GAT GTT AAG CTT GCT TTA CAG GGG TTG AAT GAT CTA TTA AAT GGG AGC AAA GCA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   D   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAN CTT GCT TTA CAG GGG TTG AAT GAT CTA TTA AAT GGG AGC AAA GCA

                   410                                               420                                               430
-------------------+-------------------------------------------------+-------------------------------------------------+---------
 K   T   S   S   D   F   S   A   W   H   N   E   L   D   Q   Q   K   R   E   F   P   L   G   Y   K   T   F
AAG ACA AGT TCT GAT TTT AGT GCA TGG CAC AAT GAG TTG GAC CAG CAG AAG AGG GAG TTT CCT CTG GGG TAC AAA ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAG AAG AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAG AAG AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAG AAG AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAG AAG AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAN AAN AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT
```

Figure 8 Continued

```
                                          440                                                   450
------------------------------------------+---------------------------------------------------+---------------------------------------------
 G   E   E   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGT GAA GAG ATC CCA CCG CAA TAT GCC ATT CAG GTG CTG GAT GAG CTG ACG AAA GGT GAG GCA ATC ATC GCT ACT GGT GTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGC GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACT GGT GTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGC GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACT GGT GTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGC GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACT GGT GTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGC GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACT GGT GTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGC GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACT GGT GTT

460                                       470                                                480
-+----------------------------------------+--------------------------------------------------+-------------------------
 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   A   G   L   G   A
GGG CAG CAC CAG ATG TGG GCG GCA CAA TAT TAC ACC TAC AAG CGG CCA CGG CAG TGG CTG TCT TCG GCT GGT CTG GGC GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAC CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCG TCT GGT TTG GGG GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAC CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCG TCT GGT TTG GGG GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAC CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCG TCT GGT TTG GGG GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAC CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCG TCT GGT TTG GGG GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAC CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCG TCT GGT TTG GGG GCA
```

Figure 8 Continued

```
            490                                                500                                                510
------------+--------------------------------------------------+--------------------------------------------------+------------
 M   G   F   G   L   P   A   A   A   G   A   S   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG CTG CCT GCT GCA GCT GGT GCT TCT GTG GCT AAC CCA GGT GTC ACA GTT GTT GAT ATT GAT GGG GAT GGT

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTA CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGT GAT GGT

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTA CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGT GAT GGT

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTA CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGT GAT GGT

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTA CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGT GAT GGT

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTA CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGT GAT GGT

                    520                                                530                                                540
--------------------+--------------------------------------------------+--------------------------------------------------+-
 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   V   L   N   N   Q
AGC TTC CTC ATG AAC ATT CAG GAG CTG GCA TTG ATC CGC ATT GAG AAC CTC CCT GTG AAG GTG ATG GTG TTG AAC AAC CAA

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGC ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGC ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGC ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGC ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGC ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG
```

Figure 8 Continued

```
                                   550                                                   560
-----------------------------------+-------------------------------------------------------+-------------------------
 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   C
CAT TTG GGT ATG GTG GTG CAA TGG GAG GAT AGG TTT TAC AAG GCG AAT AGG GCG CAT ACA TAC TTG GGC AAC CCG GAA TGT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAT CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAT CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAT CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAT CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAT CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

         570                                      580                                        590
---------+----------------------------------------+------------------------------------------+-----------------
 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   I   P   A   V   R   V   T   K   K   S   E
GAG AGC GAG ATA TAT CCA GAT TTT GTG ACT ATT GCT AAG GGG TTC AAT ATT CCT GCA GTC CGT GTA ACA AAG AAG AGT GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCA GCA GTT CGA GTG ACG AAG AAG AGC GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCA GCA GTT CGA GTG ACG AAG AAG AGC GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCA GCA GTT CGA GTG ACG AAG AAG AGC GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCA GCA GTT CGA GTG ACG AAG AAG AGC GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCA GCA GTT CGA GTG ACG AAG AAG AGC GAA
```

Figure 8 Continued

```
                     600                                    610                                    620
---------------------+--------------------------------------+--------------------------------------+-----
V   R   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC CGT GCC GCC ATC AAG AAG ATG CTC GAG ACT CCA GGG CCA TAC TTG TTG GAT ATC ATC GTC CCG CAC CAG GAG CAT GTG

V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATA GTC CCG CAT CAG GAG CAC GTG

V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATA GTC CCG CAT CAG GAG CAC GTG

V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATA GTC CCG CAT CAG GAG CAC GTG

V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATA GTC CCG CAT CAG GAG CAC GTG

V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATA GTC CCG CAT CAG GAG CAC GTG

                             630                                    640
-----------------------------+--------------------------------------+--------------------
L   P   M   I   P   S   G   G   A   F   K   D   M   I   L   D   G   D   G   R   T   V   Y   .
CTG CCT ATG ATC CCA AGT GGG GGC GCA TTC AAG GAC ATG ATC CTG GAT GGT GAT GGC AGG ACT GTG TAT TAA

L   P   M   I   P   S   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AGC GGT GGT GCT TTC AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC

L   P   M   I   P   N   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AAC GGT GGT GCT TTC AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC

L   P   M   I   P   N   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AAC GGT GGT GCT TTC AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC

L   P   M   I   P   S   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AGC GGT GGT GCT TTC AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC

L   P   M   I   P   N   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S   Y
CTG CCT ATG ATC CCA AAC GGT GGT GCT TTC AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC TGA
```

Figure 9

```
                                      10                                                 20                                    SEQ ID NO:
-----------------------  -----------+-------------------------------------------+-----------------------------
 M   A   T   T   A   A   A   A   A   A   A   L   S   A   A   A   T   A   K   T   G   R   K   N   H   Q   R      14
ATG GCT ACG ACC GCC GCG GCC GCG GCC GCC GCC CTG TCC GCC GCC GCG ACG GCC AAG ACC GGC CGT AAG AAC CAC CAG CGA     13

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      26
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     25

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      28
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     27

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      30
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     29

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      32
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     31

 V   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      34
GT- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     33

        30                                      40                                      50
---------+---------------------------------------+---------------------------------------+---------------
 H   H   V   L   P   A   R   G   R   V   G   A   A   A   V   R   C   S   A   V   S   P   V   T   P   P   S      14
CAC CAC GTC CTT CCC GCT CGA GGC CGG GTG GGG GCG GCG GCG GTC AGG TGC TCG GCG GTG TCC CCG GTC ACC CCG CCG TCC     13

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      26
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     25

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      28
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     27

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      30
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     29

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      32
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     31

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X      34
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---     33
```

Figure 9 Continued

```
                                                  60                                                          70                                                          80
---------------------------------------------------+-----------------------------------------------------------+-----------------------------------------------------------+-------
 P   A   P   P   A   T   P   L   R   P   W   G   P   A   E   P   R   K   G   A   D   I   L   V   E   A   L
CCG GCG CCG CCG GCC ACG CCG CTC CGG CCG TGG GGG CCG GCC GAG CCC CGC AAG GGC GCG GAC ATC CTC GTG GAG GCG CTG

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

 X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X   X
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---

                                90                                      100
--------------------------------+---------------------------------------+-------------------------------
 E   R   C   G   V   S   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
GAG CGG TGC GGC GTC AGC GAC GTG TTC GCC TAC CCG GGC GGC GCG TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCC CCG

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCT GGC GGC GCG TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCA

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCT GGC GGC GCG TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCA

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCT GGC GGC GCG TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCA

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCT GGC GGC GCG TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCA

 X   X   X   X   X   X   D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
--- --- --- --- --- --C GAC GTC TTC GCC TAC CCT GGC GGC GCG TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCA
```

Figure 9 Continued

```
    110                                               120                                                130
-----+-------------------------------------------------+-------------------------------------------------+-------------------
V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGC GAG GCG TTC GCG GCG TCC GGG TAC GCG CGC GCG TCC GGC CGC GTC

V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGG TAC GCC CGC GCG TCC GGC CGC GTC

V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGG TAC GCC CGC GCG TCC GGC CGC GTC

V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGG TAC GCC CGC GCG TCC GGC CGC GTC

V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGG TAC GCC CGC GCG TCC GGC CGC GTC

V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGG TAC GCC CGC GCG TCC GGC CGC GTC

                140                                               150                                             160
----------------+-------------------------------------------------+-------------------------------------------------+---------
G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   V   P
GGG GTC TGC GTC GCC ACC TCC GGC CCC GGG GCA ACC AAC CTC GTG TCC GCG CTC GCC GAC GCG CTG CTC GAC TCC GTC CCG

G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCT CTC CTC GAC TCC ATC CCC

G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCT CTC CTC GAC TCC ATC CCC

G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCT CTC CTC GAC TCC ATC CCC

G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCT CTC CTC GAC TCC ATC CCC

G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCT CTC CTC GAC TCC ATC CCC
```

Figure 9 Continued

```
                                        170                                                         180
--------------------------------------------+-----------------------------------------------------------+------------------------------------
 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACC GAC GCC TTC CAG GAG ACG CCC ATA GTC GAG GTC ACC

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAT GCG TTC CAG GAG ACG CCC ATC GTG GAG GTC ACG

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAT GCG TTC CAG GAG ACG CCC ATC GTG GAG GTC ACG

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAT GCG TTC CAG GAG ACG CCC ATC GTG GAG GTC ACG

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAT GCG TTC CAG GAG ACG CCC ATC GTG GAG GTC ACG

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAT GCG TTC CAG GAG ACG CCC ATC GTG GAG GTC ACG

190                                         200                                                         210
-+------------------------------------------+-----------------------------------------------------------+------------------------
 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAT TAC CTT GTC CTT GAT GTG GAG GAC ATC CCC CGC GTC ATA CAG GAA GCC TTC TTC CTC GCG

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTC GCA

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTC GCA

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTC GCA

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTC GCA

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTC GCA
```

Figure 9 Continued

```
                220                                     230                                     240
----------------+---------------------------------------+---------------------------------------+------------
 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCG GGC CGT CCT GGC CCG GTG CTG GTC GAC ATC CCC AAG GAC ATC CAG CAG CAG ATG GCC GTG CCG GTC TGG GAC ACC

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTG GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCT GTC TGG GAC ACG

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTG GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCT GTC TGG GAC ACG

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTG GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCT GTC TGG GAC ACG

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTG GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCT GTC TGG GAC ACG

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTG GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCT GTC TGG GAC ACG

                            250                                     260                                     270
----------------------------+---------------------------------------+---------------------------------------+-
 S   M   N   L   P   G   Y   I   A   R   L   P   K   P   P   A   T   E   L   L   E   Q   V   L   R   L   V
TCG ATG AAT CTA CCA GGG TAC ATC GCA CGC CTG CCC AAG CCA CCC GCG ACA GAA TTG CTT GAG CAG GTC TTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCG ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCG ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCG ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCG ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCG ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT
```

Figure 9 Continued

```
                                   280                                         290
-----------------------------------+-------------------------------------------+--------------------
 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   S   A   S   G   D   E   L   R   W   F   V   E   L
GGC GAG TCA CGG CGC CCG ATT CTC TAT GTC GGT GGT GGC TGC TCT GCA TCT GGT GAC GAA TTG CGC TGG TTT GTT GAG CTG

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCA TCT GGT GAG GAG TTG CGC CGC TTT GTT GAG CTC

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCA TCT GGT GAG GAG TTG CGC CGC TTT GTT GAG CTC

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCA TCT GGT GAG GAG TTG CGC CGC TTT GTT GAG CTC

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCA TCT GGT GAG GAG TTG CGC CGC TTT GTT GAG CTC

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCA TCT GGT GAG GAG TTG CGC CGC TTT GTT GAG CTC

       300                                         310                                         320
-------+-------------------------------------------+-------------------------------------------+--------
 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGT ATC CCA GTT ACA ACC ACT CTG ATG GGC CTC GGC AAT TTC CCC AGT GAC GAC CCG TTG TCC CTG CGC ATG CTT GGG

 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTT ATG GGC CTT GGC AAC TTC CC- AGT GAC GAC CCA CTG TCT CTG CGC ATG CTG GGG

 T   G   I   X   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT C-A GTT ACA ACT ACT CTT ATG GGC CTT GGC AAC TTC CCC AGT GAC GAC CCA CTG TCT CTG CGC ATG CTG GGG

 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTT ATG GGC CTT GGC AAC TTC CCC AGT GAC GAC CCA CTG TCT CTG CGC ATG CTG GGG

 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTT ATG GGC CTT GGC AAC TTC CCC AGT GAC GAC CCA CTG TCT CTG CGC ATG CTG GGG

 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTT ATG GGC CTT GGC AAC TTC CCC AGT GAC GAC CCA CTG TCT CTG CGC ATG CTG GGG
```

Figure 9 Continued

```
                           330                                     340                                     350
---------------------------+---------------------------------------+---------------------------------------+-----
 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   G   V   R   F   D   D   R
ATG CAT GGC ACG GTG TAC GCA AAT TAT GCC GTG GAT AAG GCT GAC CTG TTG CTT GCG TTT GGT GTG CGG TTT GAT GAT CGT

 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   I   A   F   G   V   R   F   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG NTT GCA TTT GGT GTG CGG TTT GAT GAT CGT

 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   G   V   R   F   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG CTT GCA TTT GGT GTG CGG TTT GAT GAT CGT

 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   G   V   Q   V   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG CTT GCA TTT GGT GTG CNG GTT GAT GAT CGT

 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   G   V   R   F   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG CTT GCA TTT GGT GTG CGG TTT GAT GAT CGT

 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   G   V   R   F   D   D   R
ATG CAT GGN ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG CTT GCA TTT GGT GTG CGG TTT GAT GAT CGT

                                    360                                     370
------------------------------------+---------------------------------------+------------------------------
 V   T   G   K   I   E   A   F   A   S   R   A   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACA GGG AAA ATT GAG GCT TTT GCA AGC AGG GCC AAG ATT GTG CAC ATT GAC ATT GAT CCA GCA GAG ATT GGA AAG AAC

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACC GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GTG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACC GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GTG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC

 V   N   G   K   I   E   A   F   A   S   R   S   K   I   E   H   I   D   I   D   P   A   E   I   G   K   N
GTG ANC GGG AAA ATC GAN GCT TTT GCA AGC AGG TCC AAG ATT GNG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACC GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GTG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACC GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GTG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC
```

Figure 9 Continued

```
     380                                         390                                         400
-----+-------------------------------------------+-------------------------------------------+---------------------
 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   A   L   L   Q   Q   S   T   T
AAG CAA CCA CAT GTG TCA ATT TGC GCA GAT GTT AAG CTT GCT TTA CAG GGC TTG AAT GCT CTG CTA CAA CAG AGC ACA ACA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   A   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAG CTT GCT TTA CAG GGG TTG AAT GCT CTA TTA AAT GGG AGC AAA GCA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   A   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAG CTT GCT TTA CAG GGG TTG AAT GCT CTA TTA AAT GGG AGC AAA GCA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   A   L   L   N   G   S   K   A
AAG CAG CCA CNT GTC TCC ATT TGT GCA GAT GTT AAN CTT GCT TTA CAG GGG TTG AAT GCN CTA TTA AAT GGG AGC AAA GCA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   A   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAG CTT GCT TTA CAG GGG TTG AAT GCT CTA TTA AAT GGG AGC AAA GCA

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   A   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAG CTT GCT TTA CAG GGG TTG AAT GCT CTA TTA AAT GGG AGC AAA GCA

                 410                                         420                                         430
-----------------+-------------------------------------------+-------------------------------------------+---------
 K   T   S   S   D   F   S   A   W   H   N   E   L   D   Q   Q   K   R   E   F   P   L   G   Y   K   T   F
AAG ACA AGT TCT GAT TTT AGT GCA TGG CAC AAT GAG TTG GAC CAG CAG AAG AGG GAG TTT CCT CTG GGG TAC AAA ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAG AAG AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAG AAG AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   K   E   F   P   L   G   F   K   T   F
CAA CAG GGN CTG GAT TTT GGT CCA TGG CNC AAG GAG TTG GAT CAG CAA AAG ANG GAG TTT CCT CTA GGA TTC AAN ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAG AAG AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAG AAG AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT
```

Figure 9 Continued

```
                                 440                                                            450
------------------------------------+---------------------------------------------------------+---------------------------------------
 G   E   E   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGT GAA GAG ATC CCA CCG CAA TAT GCC ATT CAG GTG CTG GAT GAG CTG ACG AAA GGT GAG GCA ATC ATC GCT ACT GGT GTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGT GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACC GGT GTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGT GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACC GGT GTT

 G   E   A   I   P   P   P   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGN GAN GCC ATC CCG CCG CCA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACC GGN GTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGT GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACC GGT GTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGT GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACN AAA GGG GAG GCG ATC ATT GCC ACC GGT GTT

460                                                 470                                                 480
-+--------------------------------------------------+--------------------------------------------------+-----------------------
 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   A   G   L   G   A
GGG CAG CAC CAG ATG TGG GCG GCA CAA TAT TAC ACC TAC AAG CGG CCA CGG CAG TGG CTG TCT TCG GCT GGT CTG GGC GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAT CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCA TCC GGT TTG GGT GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAT CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCA TCC GGT TTG GGT GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAT CAN ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCN CGG CAG TGG CTG TCT TCA TCC GGT TTG GGT GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAT CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCA TCC GGT TTG GGT GCA

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAT CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCA TCC GGT TTG GGT GCA
```

Figure 9 Continued

```
             490                                     500                                     510
-------------+---------------------------------------+---------------------------------------+------------
 M   G   F   G   L   P   A   A   A   G   A   S   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG CTG CCT GCT GCA GCT GGT GCT TCT GTG GCT AAC CCA GGT GTC ACA GTT GTT GAT ATT GAT GGG GAT GGT

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTG CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGG GAT GGT

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTG CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGG GAT GGT

 M   G   F   G   L   P   A   A   A   G   G   X   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTG CCA GCT GCA GCT GGC GGC -CT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGG GAT GGT

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTG CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGG GAT GGT

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTG CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGG GAT GGT

                         520                                     530                                   540
-------------------------+---------------------------------------+-------------------------------------+-
 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   V   L   N   N   Q
AGC TTC CTC ATG AAC ATT CAG GAG CTG GCA TTG ATC CGC ATT GAG AAC CTC CCT GTG AAG GTG ATG GTG TTG AAC AAC CAA

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGT ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGT ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGT ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGT ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGT ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG
```

Figure 9 Continued

```
                                                   550                                     560
------------------------------------------------+-------------------------------------+--------------------------
 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   C
CAT TTG GGT ATG GTG GTG CAA TGG GAG GAT AGG TTT TAC AAG GCG AAT AGG GCG CAT ACA TAC TTG GGC AAC CCG GAA TGT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAC CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAC CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAC CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAC CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAC CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

        570                                      580                                      590
----------+--------------------------------------+--------------------------------------+----------------
 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   I   P   A   V   R   V   T   K   K   S   E
GAG AGC GAG ATA TAT CCA GAT TTT GTG ACT ATT GCT AAG GGG TTC AAT ATT CCT GCA GTC CGT GTA ACA AAG AAG AGT GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCG GCA GTT CGT GTG ACG AAG AAG AGC GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCG GCA GTT CGT GTG ACG AAG AAG AGC GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCG GCA GTT CGT GTG ACG AAG AAG AGC GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCG GCA GTT CGT GTG ACG AAG AAG AGC GAA

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCG GCA GTT CGT GTG ACG AAG AAG AGC GAA
```

Figure 9 Continued

```
                    600                                     610                                     620
--------------------+---------------------------------------+---------------------------------------+-----
 V   R   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC CGT GCC GCC ATC AAG AAG ATG CTC GAG ACT CCA GGG CCA TAC TTG TTG GAT ATC ATC GTC CCG CAC CAG GAG CAT GTG

 V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATT GTC CCG CAT CAG GAG CAC GTG

 V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATT GTC CCG CAT CAG GAG CAC GTG

 V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATT GTC CCG CAT CAG GAG CAC GTG

 V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATT GTC CCG CAT CAG GAG CAC GTG

 V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATT GTC CCG CAT CAG GAG CAC GTG

                            630                                     640
----------------------------+---------------------------------------+--------------------
 L   P   M   I   P   S   G   G   A   F   K   D   M   I   L   D   G   D   G   R   T   V   Y   .
CTG CCT ATG ATC CCA AGT GGG GGC GCA TTC AAG GAC ATG ATC CTG GAT GGT GAT GGC AGG ACT GTG TAT TAA

 L   P   M   I   P   S   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AGC GGT GGT GCT TTT AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC

 L   P   M   I   P   S   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AGC GGT GGT GCT TTT AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC

 L   P   M   I   P   S   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AGC GGT GGT GCT TTT AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC

 L   P   M   I   P   N   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AAC GGT GGT GCT TTT AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC

 L   P   M   I   P   S   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AGC GGT GGT GCT TTT AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC
```

Figure 10

Partial DNA sequence and deduced amino acid sequence of TealIMI1 15A

```
SEQ ID NO:2           D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
SEQ ID NO:1         C GAC GTC TTC GCC TAC CCC GGC GGC GCC TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCC

V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGC TAC GCC CGC GCG TCC GGC CGC GTC

G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCC CTC CTC GAC TCC ATC CCC

M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAC GCG TTC CAG GAG ACG CCC ATA GTG GAG GTC ACG

R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTT GCA

S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTA GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCC GTC TGG GAC ACT

P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCA ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCG TCT GGC GAG GAG TTG CGC CGC TTT GTT GAG CTT

T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTG ATG GGC CTT GGC AAC TTC CCC AGC GAC GAC CCA CTG TCT CTG CGC ATG CTT GGG
```

Figure 10 continued

```
 M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   I   A   F   G   V   R   F   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG NTC GCA TTT GGT GTG CGG TTT GAT GAT CGT

 V   T   G   K   I   E   A   F   A   S   R   S   K   I   E   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACT GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GNG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC

 K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   D   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAN CTT GCT TTA CAG GGG TTG AAT GAT CTA TTA AAT GGG AGC AAA GCA

 Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAN AAN AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT

 G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGC GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACT GGT GTT

 G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAC CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCG TCT GGT TTG GGG GCA

 M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTA CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGT GAT GGT

 S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGC ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

 H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAT CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

 E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCA GCA GTT CGA GTG ACG AAG AAG AGC GAA
```

Figure 10 continued

```
 V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATA GTC CCG CAT CAG GAG CAC GTG

 L   P   M   I   P   N   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S   Y
CTG CCT ATG ATC CCA AAC GGT GGT GCT TTC AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC TGA
```

Figure 11

Partial DNA sequence and deduced amino acid sequence of TealIMI2 11A

```
SEQ ID NO:4          D   V   F   A   Y   P   G   G   A   S   M   E   I   H   Q   A   L   T   R   S   P
SEQ ID NO:3        C GAC GTC TTC GCC TAC CCT GGC GGC GCG TCC ATG GAG ATC CAC CAG GCG CTG ACG CGC TCG CCA

 V   I   T   N   H   L   F   R   H   E   Q   G   E   A   F   A   A   S   G   Y   A   R   A   S   G   R   V
GTC ATC ACC AAC CAC CTC TTC CGC CAC GAG CAG GGG GAG GCG TTC GCG GCG TCC GGG TAC GCC CGC GCG TCC GGC CGC GTC

 G   V   C   V   A   T   S   G   P   G   A   T   N   L   V   S   A   L   A   D   A   L   L   D   S   I   P
GGC GTC TGC GTC GCC ACC TCC GGC CCG GGG GCC ACC AAC CTC GTC TCC GCG CTC GCC GAC GCT CTC CTC GAC TCC ATC CCC

 M   V   A   I   T   G   Q   V   P   R   R   M   I   G   T   D   A   F   Q   E   T   P   I   V   E   V   T
ATG GTC GCC ATC ACG GGC CAG GTC CCC CGC CGC ATG ATC GGC ACG GAT GCG TTC CAG GAG ACG CCC ATC GTG GAG GTC ACG

 R   S   I   T   K   H   N   Y   L   V   L   D   V   E   D   I   P   R   V   I   Q   E   A   F   F   L   A
CGC TCC ATC ACC AAG CAC AAC TAC CTG GTC CTT GAC GTG GAG GAT ATC CCC CGC GTC ATC CAG GAA GCC TTC TTC CTC GCA

 S   S   G   R   P   G   P   V   L   V   D   I   P   K   D   I   Q   Q   Q   M   A   V   P   V   W   D   T
TCC TCT GGC CGC CCG GGG CCG GTG CTG GTT GAT ATC CCC AAG GAC ATC CAG CAG CAG ATG GCT GTG CCT GTC TGG GAC ACG

 P   M   S   L   P   G   Y   I   A   R   L   P   K   P   P   S   T   E   S   L   E   Q   V   L   R   L   V
CCG ATG AGT TTG CCA GGG TAC ATC GCC CGC CTG CCC AAG CCA CCA TCT ACT GAA TCG CTT GAG CAG GTC CTG CGT CTG GTT

 G   E   S   R   R   P   I   L   Y   V   G   G   G   C   A   A   S   G   E   E   L   R   R   F   V   E   L
GGC GAG TCA CGG CGC CCA ATT CTG TAT GTT GGT GGT GGC TGC GCT GCA TCT GGT GAG GAG TTG CGC CGC TTT GTT GAG CTC

 T   G   I   P   V   T   T   T   L   M   G   L   G   N   F   P   S   D   D   P   L   S   L   R   M   L   G
ACT GGG ATT CCA GTT ACA ACT ACT CTT ATG GGC CTT GGC AAC TTC CCC AGT GAC GAC CCA CTG TCT CTG CGC ATG CTG GGG
```

Figure 11 continued

```
M   H   G   T   V   Y   A   N   Y   A   V   D   K   A   D   L   L   L   A   F   G   V   R   F   D   D   R
ATG CAT GGC ACT GTG TAT GCA AAT TAT GCA GTA GAT AAG GCT GAC CTG TTG CTT GCA TTT GGT GTG CGG TTT GAT GAT CGT

V   T   G   K   I   E   A   F   A   S   R   S   K   I   V   H   I   D   I   D   P   A   E   I   G   K   N
GTG ACC GGG AAA ATC GAG GCT TTT GCA AGC AGG TCC AAG ATT GTG CAC ATT GAC ATT GAC CCA GCT GAG ATT GGC AAG AAC

K   Q   P   H   V   S   I   C   A   D   V   K   L   A   L   Q   G   L   N   A   L   L   N   G   S   K   A
AAG CAG CCA CAT GTC TCC ATT TGT GCA GAT GTT AAG CTT GCT TTA CAG GGG TTG AAT GCT CTA TTA AAT GGG AGC AAA GCA

Q   Q   G   L   D   F   G   P   W   H   K   E   L   D   Q   Q   K   R   E   F   P   L   G   F   K   T   F
CAA CAG GGT CTG GAT TTT GGT CCA TGG CAC AAG GAG TTG GAT CAG CAG AAG AGG GAG TTT CCT CTA GGA TTC AAG ACT TTT

G   E   A   I   P   P   Q   Y   A   I   Q   V   L   D   E   L   T   K   G   E   A   I   I   A   T   G   V
GGT GAG GCC ATC CCG CCG CAA TAT GCT ATC CAG GTA CTG GAT GAG CTG ACA AAA GGG GAG GCG ATC ATT GCC ACC GGT GTT

G   Q   H   Q   M   W   A   A   Q   Y   Y   T   Y   K   R   P   R   Q   W   L   S   S   S   G   L   G   A
GGG CAG CAT CAG ATG TGG GCG GCT CAG TAT TAC ACT TAC AAG CGG CCA CGG CAG TGG CTG TCT TCA TCC GGT TTG GGT GCA

M   G   F   G   L   P   A   A   A   G   A   A   V   A   N   P   G   V   T   V   V   D   I   D   G   D   G
ATG GGA TTT GGG TTG CCA GCT GCA GCT GGC GCT GCT GTG GCC AAC CCA GGT GTT ACA GTT GTT GAC ATT GAT GGG GAT GGT

S   F   L   M   N   I   Q   E   L   A   L   I   R   I   E   N   L   P   V   K   V   M   I   L   N   N   Q
AGT TTC CTC ATG AAC ATT CAG GAG TTG GCG TTG ATC CGT ATT GAG AAC CTC CCA GTG AAG GTG ATG ATA TTG AAC AAC CAG

H   L   G   M   V   V   Q   W   E   D   R   F   Y   K   A   N   R   A   H   T   Y   L   G   N   P   E   N
CAT CTG GGA ATG GTG GTG CAG TGG GAG GAT AGG TTT TAC AAG GCC AAC CGG GCG CAC ACA TAC CTT GGC AAC CCA GAA AAT

E   S   E   I   Y   P   D   F   V   T   I   A   K   G   F   N   V   P   A   V   R   V   T   K   K   S   E
GAG AGT GAG ATA TAT CCA GAT TTT GTG ACG ATT GCT AAA GGA TTC AAC GTT CCG GCA GTT CGT GTG ACG AAG AAG AGC GAA
```

Figure 11 continued

```
 V   T   A   A   I   K   K   M   L   E   T   P   G   P   Y   L   L   D   I   I   V   P   H   Q   E   H   V
GTC ACT GCA GCA ATC AAG AAG ATG CTT GAG ACC CCA GGG CCA TAC TTG TTG GAT ATC ATT GTC CCG CAT CAG GAG CAC GTG

 L   P   M   I   P   N   G   G   A   F   K   D   M   I   M   E   G   D   G   R   T   S
CTG CCT ATG ATC CCA AAC GGT GGT GCT TTT AAG GAC ATG ATC ATG GAG GGT GAT GGC AGG ACC TCG TAC
```

Figure 12

```
Als1_ORF_Teal  (SEQ ID NO:5)   (1) GTCTGCGTCGCCACCTCCGGCCCGGGGGCCACCAACCTCGTCTCCGCGCT
Als2_ORF_Teal  (SEQ ID NO:6)   (1) GTCTGCGTCGCCACCTCCGGCCCGGGGGCCACCAACCTCGTCTCCGCGCT
Als3_ORF_Teal  (SEQ ID NO:7)   (1) GTCTGCGTCGCCACCTCCGGCCCGGGGGCCACCAACCTCGTCTCCGCGCT
    Consensus                  (1) GTCTGCGTCGCCACCTCCGGCCCGGGGGCCACCAACCTCGTCTCCGCGCT
                                   51                                             100
Als1_ORF_Teal                 (51) CGCCGACGCCCTCCTCGACTCCATCCCCATGGTCGCCATCACGGGCCAGG
Als2_ORF_Teal                 (51) CGCCGACGCTCTCCTCGACTCCATCCCCATGGTCGCCATCACGGGCCAGG
Als3_ORF_Teal                 (51) CGCTGACGCCCTCCTCGACTCCATCCCCATGGTCGCCATCACGGGCCAGG
    Consensus                 (51) CGCCGACGCCCTCCTCGACTCCATCCCCATGGTCGCCATCACGGGCCAGG
                                   101                                            150
Als1_ORF_Teal                (101) TCCCCCGCCGCATGATCGGCACGGACGCGTTCCAGGAGACGCCCATAGTG
Als2_ORF_Teal                (101) TCCCCCGCCGCATGATCGGCACGGATGCGTTCCAGGAGACGCCCATCGTG
Als3_ORF_Teal                (101) TCCCCCGCCGCATGATCGGCACGGACGCGTTCCAGGAGACGCCCATAGTG
    Consensus                (101) TCCCCCGCCGCATGATCGGCACGGACGCGTTCCAGGAGACGCCCATAGTG
                                   151                                            200
Als1_ORF_Teal                (151) GAGGTCACGCGCTCCATCACCAAGCACAACTACCTGGTCCTTGACGTGGA
Als2_ORF_Teal                (151) GAGGTCACGCGCTCCATCACCAAGCACAACTACCTGGTCCTTGACGTGGA
Als3_ORF_Teal                (151) GAGGTCACGCGCTCCATCACCAAGCACAACTACCTGGTCCTTGACGTGGA
    Consensus                (151) GAGGTCACGCGCTCCATCACCAAGCACAACTACCTGGTCCTTGACGTGGA
                                   201                                            250
Als1_ORF_Teal                (201) GGATATCCCCCGCGTCATCCAGGAAGCCTTCTTCCTTGCATCCTCTGGCC
Als2_ORF_Teal                (201) GGATATCCCCCGCGTCATCCAGGAAGCCTTCTTCCTCGCATCCTCTGGCC
Als3_ORF_Teal                (201) GGATATCCCCCGCGTCATCCAGGAAGCCTTCTTCCTCGCGTCCTCTGGCC
    Consensus                (201) GGATATCCCCCGCGTCATCCAGGAAGCCTTCTTCCTCGCATCCTCTGGCC
                                   251                                            300
Als1_ORF_Teal                (251) GCCCGGGGCCGGTGCTAGTTGATATCCCCAAGGACATCCAGCAGCAGATG
Als2_ORF_Teal                (251) GCCCGGGGCCGGTGCTGGTTGATATCCCCAAGGACATCCAGCAGCAGATG
Als3_ORF_Teal                (251) GCCCGGGGCCGGTGCTGGTTGATATCCCCAAGGATATCCAGCAGCAGATG
    Consensus                (251) GCCCGGGGCCGGTGCTGGTTGATATCCCCAAGGACATCCAGCAGCAGATG
```

Figure 12 Continued

```
                         301                                              350
Als1_ORF_Teal    (301) GCTGTGCCCGTCTGGGACACTCCAATGAGTTTGCCAGGGTACATCGCCCG
Als2_ORF_Teal    (301) GCTGTGCCTGTCTGGGACACGCCGATGAGTTTGCCAGGGTACATCGCCCG
Als3_ORF_Teal    (301) GCCGTGCCTATCTGGGACACGCCGATGAGTTTGCCAGGGTACATCGCCCG
    Consensus    (301) GCTGTGCCTGTCTGGGACACGCCGATGAGTTTGCCAGGGTACATCGCCCG
                         351                                              400
Als1_ORF_Teal    (351) CCTGCCCAAGCCACCATCTACTGAATCGCTTGAGCAGGTCCTGCGTCTGG
Als2_ORF_Teal    (351) CCTGCCCAAGCCACCATCTACTGAATCGCTTGAGCAGGTCCTGCGTCTGG
Als3_ORF_Teal    (351) CCTGCCCAAGCCACCATCTACTGAATCGCTTGAGCAGGTCCTGCGTCTGG
    Consensus    (351) CCTGCCCAAGCCACCATCTACTGAATCGCTTGAGCAGGTCCTGCGTCTGG
                         401                                              450
Als1_ORF_Teal    (401) TTGGCGAGTCACGGCGCCCAATTCTGTATGTTGGTGGTGGCTGCGCTGCG
Als2_ORF_Teal    (401) TTGGCGAGTCACGGCGCCCAATTCTGTATGTTGGTGGTGGCTGCGCTGCA
Als3_ORF_Teal    (401) TTGGCGAGTCACGGCGCCCAATTCTGTATGTTGGTGGTGGCTGCGCTGCA
    Consensus    (401) TTGGCGAGTCACGGCGCCCAATTCTGTATGTTGGTGGTGGCTGCGCTGCA
                         451                                              500
Als1_ORF_Teal    (451) TCTGGCGAGGAGTTGCGCCGCTTTGTTGAGCTTACTGGGATTCCAGTTAC
Als2_ORF_Teal    (451) TCTGGTGAGGAGTTGCGCCGCTTTGTTGAGCTCACTGGGATTCCAGTTAC
Als3_ORF_Teal    (451) TCCGGCGAGGAGTTGCGCCGCTTTGTTGAGCTCACTGGGATTCCGGTTAC
    Consensus    (451) TCTGGCGAGGAGTTGCGCCGCTTTGTTGAGCTCACTGGGATTCCAGTTAC
                         501                                              550
Als1_ORF_Teal    (501) AACTACTCTGATGGGCCTTGGCAACTTCCCCAGCGACGACCCACTGTCTC
Als2_ORF_Teal    (501) AACTACTCTTATGGGCCTTGGCAACTTCCCCAGTGACGACCCACTGTCTC
Als3_ORF_Teal    (501) AACTACTCTGATGGGCCTTGGCAACTTCCCCAGCGACGACCCACTGTCTC
    Consensus    (501) AACTACTCTGATGGGCCTTGGCAACTTCCCCAGCGACGACCCACTGTCTC
                         551                                              600
Als1_ORF_Teal    (551) TGCGCATGCTTGGGATGCATGGCACTGTGTATGCAAATTATGCAGTAGAT
Als2_ORF_Teal    (551) TGCGCATGCTGGGGATGCATGGCACTGTGTATGCAAATTATGCAGTAGAT
Als3_ORF_Teal    (551) TGCGCATGCTTGGGATGCATGGCACTGTGTATGCAAATTATGCAGTCGAT
    Consensus    (551) TGCGCATGCTTGGGATGCATGGCACTGTGTATGCAAATTATGCAGTAGAT
```

Figure 12 Continued

```
                     601                                              650
Als1_ORF_Teal  (601) AAGGCTGACCTGTTGCTCGCATTTGGTGTGCGGTTTGATGATCGTGTGAC
Als2_ORF_Teal  (601) AAGGCTGACCTGTTGCTTGCATTTGGTGTGCGGTTTGATGATCGTGTGAC
Als3_ORF_Teal  (601) AAGGCTGACCTGTTGCTTGCATTTGGTGTGCGGTTTGATGATCGCGTGAC
    Consensus  (601) AAGGCTGACCTGTTGCTTGCATTTGGTGTGCGGTTTGATGATCGTGTGAC
                     651                                              700
Als1_ORF_Teal  (651) TGGGAAAATCGAGGCTTTTGCAAGCAGGTCCAAGATTGTGCACATTGACA
Als2_ORF_Teal  (651) CGGGAAAATCGAGGCTTTTGCAAGCAGGTCCAAGATTGTGCACATTGACA
Als3_ORF_Teal  (651) TGGGAAAATCGAGGCCTTTGCAAGCAGGTCCAAGATTGTGCACATTGACA
    Consensus  (651) TGGGAAAATCGAGGCTTTTGCAAGCAGGTCCAAGATTGTGCACATTGACA
                     701                                              750
Als1_ORF_Teal  (701) TTGACCCAGCTGAGATTGGCAAGAACAAGCAGCCACATGTCTCCATTTGT
Als2_ORF_Teal  (701) TTGACCCAGCTGAGATTGGCAAGAACAAGCAGCCACATGTCTCCATTTGT
Als3_ORF_Teal  (701) TTGACCCAGCTGAGATTGGCAAGAACAAGCAGCCACATGTCTCCATTTGT
    Consensus  (701) TTGACCCAGCTGAGATTGGCAAGAACAAGCAGCCACATGTCTCCATTTGT
                     751                                              800
Als1_ORF_Teal  (751) GCAGATGTTAAGCTTGCTTTACAGGGGTTGAATGATCTATTAAATGGGAG
Als2_ORF_Teal  (751) GCAGATGTTAAGCTTGCTTTACAGGGGTTGAATGCTCTATTAAATGGGAG
Als3_ORF_Teal  (751) GCAGATGTTAAGCTTGCTTTACAGGGGTTGAATGCTCTATTAAATGGGAG
    Consensus  (751) GCAGATGTTAAGCTTGCTTTACAGGGGTTGAATGCTCTATTAAATGGGAG
                     801                                              850
Als1_ORF_Teal  (801) CAAAGCACAACAGGGTCTGGATTTTGGTCCATGGCACAAGGAGTTGGATC
Als2_ORF_Teal  (801) CAAAGCACAACAGGGTCTGGATTTTGGTCCATGGCACAAGGAGTTGGATC
Als3_ORF_Teal  (801) CAAAGCACAACAGGGTCTGGATTTTGGTCCATGGCACAAGGAGTTGGATC
    Consensus  (801) CAAAGCACAACAGGGTCTGGATTTTGGTCCATGGCACAAGGAGTTGGATC
                     851                                              900
Als1_ORF_Teal  (851) AGCAGAAGAGGGAGTTTCCTCTAGGATTCAAGACTTTTGGCGAGGCCATC
Als2_ORF_Teal  (851) AGCAGAAGAGGGAGTTTCCTCTAGGATTCAAGACTTTTGGTGAGGCCATC
Als3_ORF_Teal  (851) AGCAGAAGAGGGAGTTTCCTCTAGGATTCAAGACTTTTGGCGAGGCCATC
    Consensus  (851) AGCAGAAGAGGGAGTTTCCTCTAGGATTCAAGACTTTTGGCGAGGCCATC
```

Figure 12 Continued

```
                          901                                               950
  Als1_ORF_Teal   (901)   CCGCCGCAATATGCTATCCAGGTACTGGATGAGCTGACAAAAGGGGAGGC
  Als2_ORF_Teal   (901)   CCGCCGCAATATGCTATCCAGGTACTGGATGAGCTGACAAAAGGGGAGGC
  Als3_ORF_Teal   (901)   CCGCCGCAATATGCTATCCAGGTACTGGATGAGCTGACAAAAGGGGAGGC
      Consensus   (901)   CCGCCGCAATATGCTATCCAGGTACTGGATGAGCTGACAAAAGGGGAGGC
                          951                                              1000
  Als1_ORF_Teal   (951)   GATCATTGCCACTGGTGTTGGGCAGCACCAGATGTGGGCGGCTCAGTATT
  Als2_ORF_Teal   (951)   GATCATTGCCACCGGTGTTGGGCAGCATCAGATGTGGGCGGCTCAGTATT
  Als3_ORF_Teal   (951)   GATCATTGCTACTGGTGTTGGGCAGCACCAGATGTGGGCGGCTCAGTATT
      Consensus   (951)   GATCATTGCCACTGGTGTTGGGCAGCACCAGATGTGGGCGGCTCAGTATT
                          1001                                             1050
  Als1_ORF_Teal  (1001)   ACACTTACAAGCGGCCACGGCAGTGGCTGTCTTCGTCTGGTTTGGGGGCA
  Als2_ORF_Teal  (1001)   ACACTTACAAGCGGCCACGGCAGTGGCTGTCTTCATCCGGTTTGGGTGCA
  Als3_ORF_Teal  (1001)   ACACTTACAAGCGGCCACGGCAGTGGCTGTCTTCGTCTGGTTTGGGGGCA
      Consensus  (1001)   ACACTTACAAGCGGCCACGGCAGTGGCTGTCTTCGTCTGGTTTGGGGGCA
                          1051                                             1100
  Als1_ORF_Teal  (1051)   ATGGGATTTGGGTTACCAGCTGCAGCTGGCGCTGCTGTGGCCAACCCAGG
  Als2_ORF_Teal  (1051)   ATGGGATTTGGGTTGCCAGCTGCAGCTGGCGCTGCTGTGGCCAACCCAGG
  Als3_ORF_Teal  (1051)   ATGGGATTTGGGTTACCAGCTGCAGCTGGCGCTGCTGTGGCCAACCCAGG
      Consensus  (1051)   ATGGGATTTGGGTTACCAGCTGCAGCTGGCGCTGCTGTGGCCAACCCAGG
                          1101                                             1150
  Als1_ORF_Teal  (1101)   TGTTACAGTTGTTGACATTGATGGTGATGGTAGTTTCCTCATGAACATTC
  Als2_ORF_Teal  (1101)   TGTTACAGTTGTTGACATTGATGGGGATGGTAGTTTCCTCATGAACATTC
  Als3_ORF_Teal  (1101)   TGTTACAGTTGTTGACATTGATGGAGATGGTAGTTTCCTCATGAACATTC
      Consensus  (1101)   TGTTACAGTTGTTGACATTGATGG GATGGTAGTTTCCTCATGAACATTC
                          1151                                             1200
  Als1_ORF_Teal  (1151)   AGGAGTTGGCGTTGATCCGCATTGAGAACCTCCCAGTGAAGGTGATGATA
  Als2_ORF_Teal  (1151)   AGGAGTTGGCGTTGATCCGTATTGAGAACCTCCCAGTGAAGGTGATGATA
  Als3_ORF_Teal  (1151)   AGGAGTTGGCATTGATCCGTATTGAGAACCTCCCTGTGAAGGTGATGATA
      Consensus  (1151)   AGGAGTTGGCGTTGATCCGTATTGAGAACCTCCCAGTGAAGGTGATGATA
```

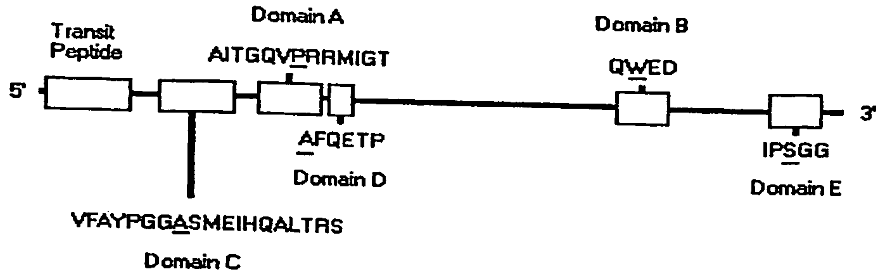
Figure 12 Continued

```
                          1201                                              1250
Als1_ORF_Teal    (1201)  TTGAACAACCAGCATCTGGGAATGGTGGTGCAGTGGGAGGATAGGTTTTA
Als2_ORF_Teal    (1201)  TTGAACAACCAGCATCTGGGAATGGTGGTGCAGTGGGAGGATAGGTTTTA
Als3_ORF_Teal    (1201)  TTGAACAACCAGCATCTGGGAATGGTGGTGCAATGGGAGGATAGGTTTTA
    Consensus    (1201)  TTGAACAACCAGCATCTGGGAATGGTGGTGCAGTGGGAGGATAGGTTTTA
                          1251                                              1300
Als1_ORF_Teal    (1251)  CAAGGCCAATCGGGCGCACACATACCTTGGCAACCCAGAAAATGAGAGTG
Als2_ORF_Teal    (1251)  CAAGGCCAACCGGGCGCACACATACCTTGGCAACCCAGAAAATGAGAGTG
Als3_ORF_Teal    (1251)  CAAGGCCAATCGGGCGCACACATACCTTGGCAACCCAGAAAATGAGAGTG
    Consensus    (1251)  CAAGGCCAATCGGGCGCACACATACCTTGGCAACCCAGAAAATGAGAGTG
                          1301                                              1350
Als1_ORF_Teal    (1301)  AGATATATCCAGATTTTGTGACGATTGCTAAAGGATTCAACGTTCCAGCA
Als2_ORF_Teal    (1301)  AGATATATCCAGATTTTGTGACGATTGCTAAAGGATTCAACGTTCCGGCA
Als3_ORF_Teal    (1301)  AGATATATCCAGATTTTGTGACGATTGCTAAAGGATTCAACGTTCCGGCA
    Consensus    (1301)  AGATATATCCAGATTTTGTGACGATTGCTAAAGGATTCAACGTTCCGGCA
                          1351                                              1400
Als1_ORF_Teal    (1351)  GTTCGAGTGACGAAGAAGAGCGAAGTCACTGCAGCAATCAAGAAGATGCT
Als2_ORF_Teal    (1351)  GTTCGTGTGACGAAGAAGAGCGAAGTCACTGCAGCAATCAAGAAGATGCT
Als3_ORF_Teal    (1351)  GTTCGTGTGACGAAGAAGAGCGAAGTCACTGCAGCAATCAAGAAGATGCT
    Consensus    (1351)  GTTCGTGTGACGAAGAAGAGCGAAGTCACTGCAGCAATCAAGAAGATGCT
                          1401                                              1450
Als1_ORF_Teal    (1401)  TGAGACCCCAGGGCCATACTTGTTGGATATCATAGTCCCGCATCAGGAGC
Als2_ORF_Teal    (1401)  TGAGACCCCAGGGCCATACTTGTTGGATATCATTGTCCCGCATCAGGAGC
Als3_ORF_Teal    (1401)  TGAGACCCCAGGGCCATACTTGTTGGATATCATCGTCCCGCATCAGGAGC
    Consensus    (1401)  TGAGACCCCAGGGCCATACTTGTTGGATATCAT GTCCCGCATCAGGAGC
                          1451                                              1500
Als1_ORF_Teal    (1451)  ACGTGCTGCCTATGATCCCAAGCGGTGGTGCTTTCAAGGACATGATCATG
Als2_ORF_Teal    (1451)  ACGTGCTGCCTATGATCCCAAGCGGTGGTGCTTTTAAGGACATGATCATG
Als3_ORF_Teal    (1451)  ACGTGCTGCCTATGATCCCAAGCGGTGGTGCTTTCAAGGACATGATCATG
    Consensus    (1451)  ACGTGCTGCCTATGATCCCAAGCGGTGGTGCTTTCAAGGACATGATCATG
```

Figure 12 Continued

```
                          1501                    1524
Als1_ORF_Teal     (1501)  GAGGGTGATGGCAGGACCTCGTAC
Als2_ORF_Teal     (1501)  GAGGGTGATGGCAGGACCTCGTAC
Als3_ORF_Teal     (1501)  GAGGGTGATGGCAGGACCTCGTAC
    Consensus     (1501)  GAGGGTGATGGCAGGACCTCGTAC
```

Figure 13


Transit Peptide
Domain A
AITGQVPRRMIGT
Domain B
QWED
5'
3'
AFQETP
Domain D
IPSGG
Domain E
VFAYPGGASMEIHQALTRS
Domain C

INTERNATIONAL SEARCH REPORT

Inter   ›nal Application No
PCT/CA 02/01051

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/60 C12N15/82 C12N9/88 C12N5/10 A01H5/00
A01H5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, EMBASE, SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 928 937 A (STOCKTON GERALD W ET AL) 27 July 1999 (1999-07-27)	44, 62-64, 70-72
Y	SEQ ID NO:3 is 92.4% identical to SEQ ID NO:4 and 91.7% identical to SEQ ID NO:2, more than 60 consecutive identical nucleotides the whole document ---	6,53,56, 61,64,69
X	EP 0 525 384 A (AMERICAN CYANAMID CO) 3 February 1993 (1993-02-03)	44, 62-64, 70-72
Y	the whole document --- -/--	6,53,56, 61,64,69

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 December 2002	02/01/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Kania, T

Form PCT/ISA/210 (second sheet) (July 1992)

Inte   onal Application No
PCT/CA 02/01051

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	NEWHOUSE K E ET AL: "TOLERANCE TO IMIDAZOLINONE HERBICIDES IN WHEAT" PLANT PHYSIOLOGY, AMERICAN SOCIETY OF PLANT PHYSIOLOGISTS, ROCKVILLE, MD, US, vol. 100, no. 2, 16 March 1992 (1992-03-16), pages 882-886, XP001093916 ISSN: 0032-0889 cited in the application the whole document ---	1-5, 9-11, 13-15, 20-28, 31-39, 49-52, 54,55, 57-60, 62,63, 65-68
Y	EP 0 508 161 A (AMERICAN CYANAMID CO) 14 October 1992 (1992-10-14) the whole document ---	1-5, 9-11, 13-15, 20-28, 31-39, 49-52, 54,55, 57-60, 62,63, 65-68
Y	SWANSON E B ET AL: "MICROSPORE MUTAGENESIS AND SELECTION CANOLA PLANTS WITH FIELD TOLERANCE TO THE IMIDAZOLINONES" THEORETICAL AND APPLIED GENETICS, vol. 78, no. 4, 1989, pages 525-530, XP009002483 ISSN: 0040-5752 cited in the application the whole document ---	1-5, 9-11, 13-15, 20-28, 31-39, 49-69
A	CHONG CHOM-KYU ET AL: "Amino acid residues conferring herbicide tolerance in tobacco acetolactate synthase." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 279, no. 2, 20 December 2000 (2000-12-20), pages 462-467, XP002224077 ISSN: 0006-291X the whole document ---	3-5
A	WO 90 14000 A (ICI PLC) 29 November 1990 (1990-11-29) ---	
A	US 5 013 659 A (CHALEFF ROY S ET AL) 7 May 1991 (1991-05-07) ---	

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	WRIGHT T R ET AL: "Cell selection and inheritance of imidazolinone resistance in sugarbeet (Beta vulgaris)." THEORETICAL AND APPLIED GENETICS, vol. 96, no. 5, April 1998 (1998-04), pages 612-620, XP002224078 ISSN: 0040-5752 page 618 -page 619 abstract	
A	SHANER DALE L ET AL: "Imidazolinone-resistant crops: Selection, characterization, and management." 1996 , HERBICIDE-RESISTANT CROPS: AGRICULTURAL, ENVIRONMENTAL, ECONOMIC,, PAGE(S) 143-157 , 1996 CRC PRESS, INC.;CRC PRESS BOCA RATON, FLORIDA, USA; LONDON, ENGLAND, UK XP001118963 ISBN: 1-56670-045-0 page 147 -page 152	

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-6,9,13-27,31-39,44,49,50,53-55,57,58,61-63,65,66, 69-71 partially; 7,10,12,45,47,52,60,68 completely

   the claimed subject-matter as far as it relates to the wheat mutant PTA3955 and SEQ ID NO:1,2

2. Claims: 27-39,44,46,48,54-56,62-64, 70-72 partially; 40-43 completely

   the claimed subject-matter as far as it relates to the wheat mutant PTA3953 comprising SEQ ID NO:3,4

3. Claims: 1-7,9,13-39,44,46,48-50,53-58,61-66, 69-72 partially; 8,11,12,51,59,67

   the claimed subject-matter as far as it relates to the wheat mutant PTA3954 comprising SEQ ID NO:3,4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA 02/01051

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/CA 02/01051

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
US 5928937	A	27-07-1999	US	5853973	A	29-12-1998
			AU	5575896	A	07-11-1996
			BR	9604993	A	30-11-1999
			CA	2218526	A1	24-10-1996
			CZ	9703317	A3	17-06-1998
			EP	0821729	A1	04-02-1998
			HU	9900852	A2	28-07-1999
			JP	11504213	T	20-04-1999
			NO	974803	A	19-12-1997
			NZ	307012	A	28-01-2000
			PL	322899	A1	02-03-1998
			WO	9633270	A1	24-10-1996
			NZ	500281	A	28-02-2000
			NZ	502491	A	28-04-2000
EP 0525384	A	03-02-1993	US	5731180	A	24-03-1998
			AT	197475	T	11-11-2000
			AU	2069492	A	04-02-1993
			AU	4335296	A	02-05-1996
			AU	9046098	A	14-01-1999
			BG	61276	B1	30-04-1997
			BG	96699	A	30-06-1994
			BR	9202950	A	30-03-1993
			CA	2074854	A1	01-02-1993
			CZ	9202385	A3	16-03-1994
			DE	69231551	D1	14-12-2000
			DE	69231551	T2	15-03-2001
			DK	525384	T3	27-11-2000
			EP	0525384	A2	03-02-1993
			ES	2153352	T3	01-03-2001
			FI	923444	A	01-02-1993
			GR	3035395	T3	31-05-2001
			HU	65483	A2	28-06-1994
			IE	922504	A1	10-02-1993
			JP	5227964	A	07-09-1993
			KR	243996	B1	01-02-2000
			MX	9204420	A1	01-02-1993
			NO	923017	A	01-02-1993
			NZ	243693	A	25-06-1993
			PL	295470	A1	31-05-1993
			PT	525384	T	30-04-2001
			RO	114348	B1	30-03-1999
			SK	238592	A3	11-07-1995
			US	6444875	B1	03-09-2002
			US	5767361	A	16-06-1998
			ZA	9205734	A	28-04-1993
EP 0508161	A	14-10-1992	AT	192625	T	15-05-2000
			AU	659142	B2	11-05-1995
			AU	1471092	A	15-10-1992
			CA	2065189	A1	09-10-1992
			CN	1065573	A ,B	28-10-1992
			CS	9201052	A3	18-11-1992
			DE	69231015	D1	15-06-2000
			DE	69231015	T2	28-09-2000
			DK	508161	T3	14-08-2000
			EP	0508161	A1	14-10-1992
			EP	0965265	A2	22-12-1999

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/CA 02/01051

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
EP 0508161	A		ES	2146202	T3	01-08-2000
			FI	921528	A	09-10-1992
			GR	3034169	T3	30-11-2000
			HU	62137	A2	28-04-1993
			IE	921101	A1	21-10-1992
			JP	5219848	A	31-08-1993
			MX	9201593	A1	01-10-1992
			NZ	242228	A	25-03-1994
			PT	508161	T	31-10-2000
			RU	2091010	C1	27-09-1997
			US	6339184	B1	15-01-2002
			ZA	9202511	A	30-12-1992
WO 9014000	A	29-11-1990	AT	132000	T	15-01-1996
			AU	5632790	A	18-12-1990
			BR	9007379	A	28-04-1992
			DE	69024491	D1	08-02-1996
			DE	69024491	T2	15-05-1996
			EP	0472573	A1	04-03-1992
			ES	2080826	T3	16-02-1996
			WO	9014000	A1	29-11-1990
			HU	60589	A2	28-10-1992
			JP	4507340	T	24-12-1992
			RO	113788	B1	30-11-1998
			RU	2128426	C1	10-04-1999
US 5013659	A	07-05-1991	US	5605011	A	25-02-1997
			US	5141870	A	25-08-1992